Model Based Detection and Location of Houses as Topographic Control Points in Digital Images

Wolfgang Förstner
Institute of Photogrammetry
Keplerstr. 11, D-7000 Stuttgart 1, FRG

Abstract:

The paper describes a concept for model based location of house roofs as topographic control points in digital or digitized images. The control points are described by hand drawn orthographic sketches. The paper deals with the interpretation of the sketches, the extraction of straight line segments from the image and the use of the uncertainty of model and image features within the matching procedure.

0. Introduction

Control points play a central role for absolute orientation of aerial images, in aerial triangulation and for rectification of satellite imagery. They form the link between the geodetic reference system and the measurements taken from images and essentially are needed for the indirect determination of the orientation parameters of the sensor platform during the flight mission as long as these parameters cannot reliably be determined directly, e. g. using the GPS-Navstar system or inertial systems.

The identification of control points is usually performed by a human operator taking advantage of his interpretation capability especially if natural control points, such as buildings or roads are used. In the course of automating image analysis tasks using digital image processing techniques also the detection and location of control points has to be automated. While this demand is obvious for the rectification of satellite imagery, where control point identification still is a burden, automatic control point identification could also be advantageous for the absolute orientation of aerial images in order to avoid aerial triangulation. This of course is only of practical value if natural control points are available.

This was the motivation for the survey department (Landesvermessungsamt) at Bonn to establish a control point data base for their orthophoto production, which aims at a periodic update of the orthophoto maps every five years. The data base consists of more than 20 000 natural control points, mainly being pairs of gable points of house roofs with their X-, Y- and Z-coordinates and a description in the form of a sketch of the roof in orthographic projection. The density is high enough that in the largest part of the area for each photo (scale appr. 1 : 12 000) at least 6 to 8 points are available for absolute orientation. Together with the stored Digital Elevation Model this information is sufficient for orthophoto production without requiring an aerial triangulation.

The task of control point identification and location is a well defined one in this case, which was the reason to start an investigation how the task could be automated using techniques from image processing, pattern recognition and artificial intelligence. The task is similar to object location in robotics in case a CAD-model is available to identify objects on a conveyor belt with a digital camera, a tactile sensor or a laser ranger (cf. GRIMSON/LOZANO-PEREZ 1984, 1987; FAUGERAS/HEBERT 1987, HORAUD 1987). CAD-models as well as the models available in the mentioned control point data base are specific models, fixed or at least known up to a few parameters, e. g. the slope or the width of the roof planes. Also orientation pa-

rameters may be unknown. Thus a few numerical parameters may have to be determined or estimated during the identification and location of the objects.

The task thus differs significantly from automatic mapping. Here generic models are necessary, which do not only allow numerical parameters to vary but also relations between the geometric primitives to vary according to some rules. Such a generic model may e. g. describe all closed polygons with rectangles, where except for the orientation and the lengths of the edges also the number of the edges may be arbitrary. An even more complex generic model for buildings has been developed by FUA and HANSON (1987).

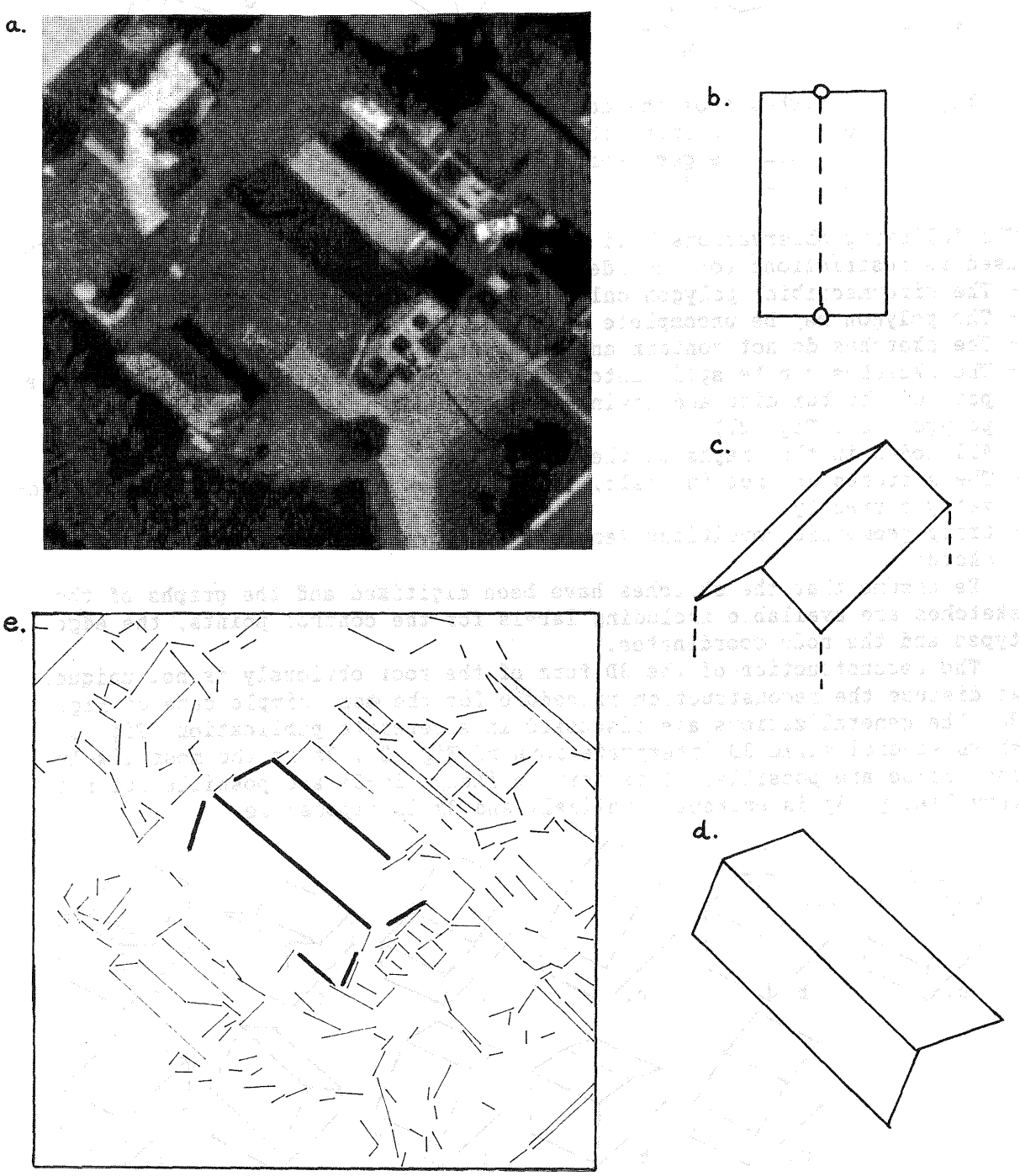
An example wants to show the principle of our approach. Fig. 1a shows a digitized subsection of an aerial image containing the house sketched in Fig. 1b. The two gable points are given with their 3D coordinates and are the actually used control point pair. Fig. 1c and 1d show the 3D-model and the model projected into the image. This projected model is to be found among the extracted straight edges shown in Fig. 1e. The bold lines indicate the solution of the match from which by a fit of the model to the image edges the position of the two gable points can be derived. The aim in this preliminary test on the feasibility of an automatic control point location procedure was to find an optimal solution in the sense that for all model edges the longest image edge had to be found. The test criteria for selecting the matches were not taking the geometric properties and the uncertainty of the edges into account. This resulted in clear deficiencies of the procedure, as the effect of the chosen thresholds onto the result could not be predicted and the quality of the result, i. e. the resultant coordinates could not be evaluated before the final decision using an intuitive measure. This experience, together with the successful identification of the correct image edges was the motivation for the more thorough statistical setup discussed in this paper.

The paper deals with the identification of 3D objects in images for which a specific model is available. As the concept has been developed on the background of the mentioned task of control point location it was based on the following line of thought:

- The task is to locate the object, i. e. determine its position, possibly its orientation in the image or in object space. The task is not to recognize the object among a given set of objects, though it can be adapted to this task.
- The model consists of a list of edges, including single points as a special case. This list needs not completely describe the object. The edges are represented by point pairs.
- The model may be uncertain. The type and degree of uncertainty has to be given explicitely or derivable from the geometry of the model.
- The model may contain symmetries or other crisp geometric conditions, such as angles being 0 ° (parallel) or 90 °.
- In case the model is only given by a 2D projection the 3D structure of the model has to be derived automatically from the projection, in general leading to a set of valid 3D-models. Heuristic conditions have to be used to reduce the ambiguity and for a ranking of the 3D-models according to their likelihood.
- The model edges are to be matched with edges extracted from the image using a heuristic search procedure.
- The uncertainty of the image edges has to be derived from the uncertainty of the intensity values taking the applied extraction procedure into account.
- The location procedure should contain selfdiagnosis capability. Thus the quality of the determined position and/or orientation has to be evaluated and be given in an intuitive form.

The paper gives an outline of the concept for solving this task. In detail the following steps are discussed: 1. Interpretation of the 2D sketches, 2. Geometric 3D-models and their projection into the image, 3. Extraction of straight edges, 4. Representation of uncertainty, 5. Matching criteria, and 6. Matching procedure.

Fig. 1 Example for matching roof model to image edges
a. digitized section of aerial image, 5 x 5 mm², 240 x 240 pixels
b. sketch of roof in control point data base
c. 3D-model of roof
d. projected model in image plane (not in scale)
e. extracted straight edges, matched edges: bold lines



1. Interpretation of the 2D sketches

Interpretation of line drawings is a classical task in artificial intelligence (cf. e. g. the overview in BRADY (1981)). Our problem is much simpler than labeling a general view of a polyhedra, as we have more a priori knowledge, specifically the known projection of the sketches and the classification of the lines within the sketches.

Fig. 2 shows 4 representative sketches from the control point data base.

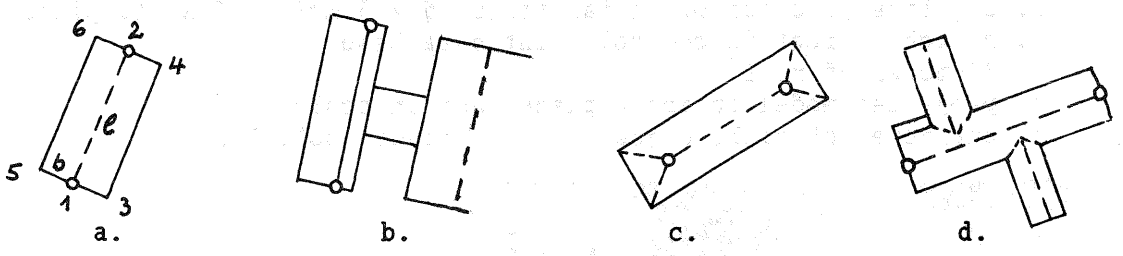


Fig. 2 Sketches from the control point data base
o points with known ground coordinates
——— edges representing vertical walls
– – – – " " intersections of roofs

The following observations hold for nearly all control points and thus are used as restrictions for the identification process:

- The circumscribing polygon only contains rectangles.
- The polygon may be uncomplete (cf. Fig. 2b).
- The sketches do not contain any information on the slope of the roofs.
- The sketches can be split into nonoverlapping basic units representing a part of the building and having no full lines inside the circumscribing polygon (cf. Fig. 2b)
- All nodes in the graphs of the basic units have order 2 or 3.
- The sketches are not in scale, though approximately representing the geometry correctly.
- Crisp geometric conditions (angles, symmetry) can be derived from the sketch.

We assume that the sketches have been digitized and the graphs of the sketches are available including labels for the control points, the edge types and the node coordinates.

The reconstruction of the 3D form of the roof obviously is not unique. We discuss the reconstruction procedure for the most simple case of Fig. 2a. The generalizations are discussed in a separate publication. Fig. 3 shows several valid 3D interpretations of Fig. 2a. 3a is the most likely one, 3b-3e are possible, 3f is very unlikely, 3g-3i are possible but not very likely, 3j is extremely unlikely and 3k is impossible.

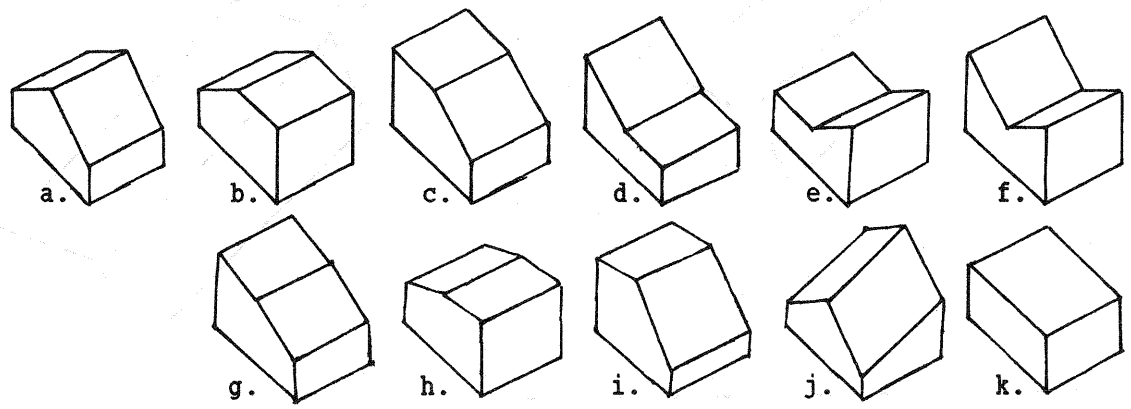
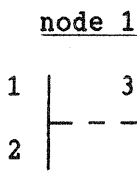


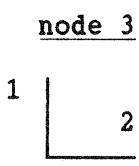
Fig. 3 3D interpretations of Fig. 2a (not in scale)

The different interpretations of the roof obviously correspond to different interpretations of the nodes and the edges resp. Specifically in this case the roofs a to h directly correspond to the labeling of node 1 in Fig. 4b. There also an example for empirically derived probabilities is given for node 1 and 3 which may have been obtained from a training set. Only the sign of the slope of the edges, as seen from the node is used in this representation.



	a	b	c	d	e	f	g	h
1	-	-	o	+	o	+	+	-
2	-	o	-	o	+	+	-	+
3	o	o	o	o	o	o	o	o

2:	+	o	-
1:+	0	.05	0
o	.05	-	.05
-	0	.05	.8



2:	+	o	-
1:+	0	.15	0
o	.15	.4	.15
-	0	.15	0

a. b. c.

Fig. 4 a. Node representation
b. Interpretations of the edges at node 1
(o = horizontal, + = ascending, - = descending)
c. Example of empirical probabilities for edge labels
(for node 1: 3 = o. The cases 3 = o and 3 = - contain 0-probabilities, - = invalid labeling)

The interpretation now can be achieved by labeling the edges at the nodes, starting with that node which leads to the most certain decision. This labeling actually is a depth first search in an interpretation tree using heuristic information to guide the search (cf. GRIMSON/ LOZANO-PEREZ 1987).

The certainty can be measured by the deviation of the probability distribution from an equal distribution, being the worst case. This deviation is the relative redundancy $R(\underline{x})$ of the decision $\underline{x}$

$$R(\underline{x}) = (H_{max} - H(\underline{x}))/H_{max} = 1 - H(\underline{x})/H_{max}$$

where $H(\underline{x})$ is the entropy of the decision, i. e. the expectation of the information $I(\underline{x}=x) = -\ln(p(\underline{x}=x))$ obtained when being told that the labeling is $\underline{x} = x$. H_{max} is the upper bound on H.

$$H(\underline{x}) = \sum_{i=1}^{n} -p(x_i)\cdot\ln(p(x_i)), \quad H_{max} = \ln(n)$$

n is the number of alternatives of the decision. Here $R_1 = R_2 = 1 - (-4\cdot 0.05\cdot\ln(0.05)-0.8\cdot\ln(0.8))/\ln(27) = 0.764$ and $R_3 = R_4 = R_5 = R_6 = 1 - (-4\cdot 0.15\cdot\ln(0.15)-0.4\cdot\ln(0.4))/\ln(9) = 0.315$. Therefore node 1 (or 2) should be labeled first, leading to the most probable interpretation (-,-,o). This decision can now be propagated until all nodes have been processed, either using geometric constraints (especially parallelism of lines belonging to one plane) or further heuristics, as used in the first decision. In this case conditional redundancies can be used as parts of the edges belonging to one node already are labeled. The likelihood of the total interpretation

then can be approximated by the negative information content $L = -\Sigma\ I(x_j) = \Sigma\ \ln(p_j)$ of the free decisions at the nodes, which needed heuristics, neglecting mutual dependencies between these decisions. This likelihood would be $L_a = \ln(0.8) = -.223$ and $L_b = \ln(0.05) = -2.996$ for the interpretations 2a and 2b resp., as the first decision, namely for labeling node 1, was uncertain while the labeling of the other nodes could use geometric constraints. This shows interpretation 2a to be more likely than interpretation 2b, as could be expected. Putting a lower bound on the likelihood of the valid interpretations or onto the number of interpretations leads to a sorted list of interpretations which then can be used as a basis for the localization.

2. Geometric 3D-models and their projection into the image

The interpretations discussed so far only give symbolic descriptions of the roofs and do not contain any geometric measures. The planimetric coordinates of the nodes can easily obtained from the sketch coordinates and the given X- and Y-coordinates of the two control points. As the height of the control points are known only the height of the remaining nodes have to be determined. This can be achieved in a simple manner:

1. Collect all nodes connected by horizontal edges with those already assigned a height to. Give these nodes the same height as those of the nodes with which they are connected to.
2. If nodes are left, which do not have a height, choose one of them which is linked to a node with a height assigned to it, choose a certain height difference for the edge, calculate the height of this new node and goto 1, else stop.

In case the graph is connected all nodes will obtain a unique height. Moreover the procedure also yields the type of link between the heights and therefore the type of geometric condition: crisp for node pairs with the same height, weak for node pairs with different heights.

As the height differences or the slopes are not given, height differences or slopes are treated as random variables with a large standard deviation which will be used to derive the covariance matrix of the 3D-model (cf. section 4). Symmetries can easily be applied in this procedure, in general reducing the number of height differences to be chosen is step 2.

The 3D-model is now projected into the image. This is possible as approximate values for the orientation parameters of the image are known from the flight plan. The X_o- and Y_o-coordinates of the projection centre are accurate up to 1 cm in the image. The flying height Z_o is known with an accuracy of about 10 % and the orientation angles can be assumed to be known with a standard deviation of a few degrees. With these approximate values the 3D-model can be projected into the image plane. Taking the uncertainty of the orientation parameters Z_o, Ω, Φ anf K into account also the covariance matrix of the projected model can be derived by error propagation. The uncertainty of the X_o- and the Y_o-coordinate is too large to be represented in the covariance matrix. Therefore the position of the projected model is assumed to be completely free within the matching process. We only assume that the house can be found in the 1 cm-neighbourhood of the projected model, requiring an area of appr. 2 x 2 cm^2 to be digitized. The result of this step thus is the projected model consisting of a list of model edges in the image coordinate system together with the covariance matrix of the coordinates of the endpoints.

3. Extracting straight edges

There exist several techniques for estracting straight edges from digital images (cf. e. g. BURNS et. al. 1986, NEVATIA/BABU 1980). We apply a different technique which is theoretically related to the operator for extracting distinct points proposed by FÖRSTNER (1986, cf. FÖRSTNER/GÜLCH 1987). The choice of the feature extraction algorithm should not have a too high impact onto the result but provide information on the quality of the extracted edges. All schemes for straight line extraction known to the author may provide this type of selfdiagnosis though they not all actually have exploited this possibility.

Our algorithm consists of the following steps (cf. FÖRSTNER 1988):

1. Determination of the normal equation matrix N for determining the optimal edge position from the following equation system $N\,\hat{y} = h$

$$\begin{bmatrix} \Sigma\, g_r^2 & \Sigma\, g_r\, g_c \\ \Sigma\, g_r\, g_c & \Sigma\, g_c^2 \end{bmatrix} \begin{bmatrix} \hat{r} \\ \hat{c} \end{bmatrix} = \begin{bmatrix} \Sigma\, r^2\, g_r^2 + \Sigma\, r\, c\, g_r\, g_c \\ \Sigma\, r\, c\, g_r\, g_c + \Sigma\, c^2\, g_c^2 \end{bmatrix}$$

   where the sums are taken over a small window, usually 3 x 3 pixels.
2. Determination of the strength $s = \mathrm{tr}N$ and the likelihood $q = 4 \cdot \det N / (\mathrm{tr}N^2)$ of the edge. q lies in the range between 0 and 1, q = 0 representing a 100 % anisotropic texture, thus suggesting an edge.
3. Extraction of edge regions given by all pixels with $s > t_s$ and $q < t_q$. The threshold t_s can be related to the noise of the intensity values g. A reaonable threshold for q has proven to lie in the range between 0.5 and 0.7.
4. Extraction of edge pixels by determining the local relative maxima in row or column direction.
5. Determination of the edge elements (edgels) using the neighbourhood of the edge pixels. One point of the straight line of the edgels can be determined from the above mentioned equation system, whereas the direction Φ can be determined from $\frac{1}{2} \cdot \arctan(2 \cdot N_{12} / (N_{11} - N_{22}))$. The position of the edgel then is the point on this line lying nearest to the centre pixel of the window. Thus the edgel is represented by its subpixel position, its orientation Φ, its strength s and its likelihood q.
6. Merging the edgels to straight edges recursively by a "line-growing" process on the 8-connected components of the edgels. The straight edge is represented by the first two moments m and M of the used edgels and the circumscribing rectangle. Starting with the edgel with the highest

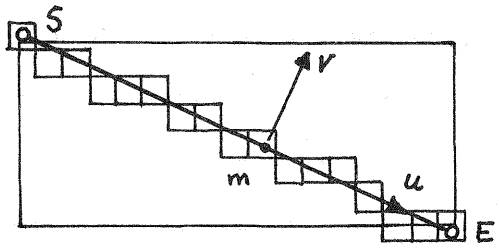


Fig. 5 Representation of edge and local coordinate system (u,v)

strength a neighbouring edgel is merged into the straight line if:

   a. its orientation Φ is close enough to the orientation α of the actual edge (e. g. $|\alpha - \Phi| < 45°$).
   b. its distance d to the actual edge is small enough (e. g. $|d| < 1$ pixel) and
   c. the smaller eigenvalue μ_2 of M, representing the strength of the actual edge, will not exceed a given threshold (e. g. $\mu_2 < 1$ pixel^2).

The final edge then is described by its endpoints which are the two points on the straight line closest to the corners of the circumscribing rectangle.

Remarks:

a. Extraction of edgels and the connection scheme are independent and can be replaced by any other algorithm.
b. A hysteresis type of thresholding (cf. CANNY 1986) can easily be integrated into the line growing process.
c. The use of subpixel estimates for the position of the edges significantly improves the resulting edges, especially the short ones.
d. In case the signal to noise ratio of the image is low an information preserving filter should be applied to suppress noise while keeping edges and corners (cf. FÖRSTNER 1988).

4. Representation of Uncertainty

In this application both, model and image edges, are uncertain due to inevitable errors in the generating process. It is crucial for the matching process to exploit this uncertainty in a consistent manner especially to achieve results which are independent of the sequence of the path in the tree search and to reduce decision errors due to inaccurate thresholds in the tests. The most appropriate way to achieve this goal is to treat the measurements as realizations of random variables and use the second moments of their distribution as representation for their uncertainty. This gives rise to simple ways for keeping track of the uncertainty by error propagation and to statistical tests of the equality of the geometric descriptions if an assumption on the type of distribution is made. This in our case will be the Gaussian because of both the central limit theorem as the principle of maximum entropy. We also can assume that within the matching process the approximate values are precise enough to be able to neglect the effects of the nonlinear relationships and use the linear surrogate models for the tests. The same arguments have been used by SMITH et. al. (1987a, b) for their "stochastic map" describing the uncertainty of the positions and orientations in robot path planning.

4.1 The uncertainty of the model

The uncertainty of the model depends on

a. the accuracy of the sketch,
b. the uncertainty of the slope(s) of the roof(parts)
c. possibly on the validity of geometric constraints derived from the sketch
d. the accuracy of the approximate values for the orientation parameters of the sensor platform.

The uncertainty of the projected model is represented by the covariance matrix of the coordinates of the nodes.

An example wants to demonstrate how this covariance matrix can be derived. Let us assume that

a. the interpretation yields the information that the roof shown in Fig. 2a consists of two rectangles in symmetric position.
b. the length-ratios between the width b and the length l of one of the rectangles have a standard deviation of 10 %,
c. the assumed height difference in the roof be h with a standard deviation of 30 %.

Starting from the given sketch points $P_1(0,0,0)$ and $P_2(0,1,0)$ the points P_3 to P_6 are given by

$$P_3(\underline{b},\ 0,\ -\underline{h}),\ P_4(\underline{b},\ 1,\ -\underline{h}),\ P_5(-\underline{b},\ 0,\ -\underline{h}),\ P_6(-\underline{b},\ 1,\ -\underline{h})$$

where stochastical variables are underlined, namely b and h having standard deviation $\sigma_b = 0.1\ b$ and $\sigma_h = 0.3\ h$ resp., b and h being the approximate values for $\underline{b}$ and $\underline{h}$ resp. As only two random variables are involved the rank of the 18 x 18 covariance matrix for the 18 coordinates of the 6 points of the 3D sketch model has rank 2 and is very sparse. Observe, that the heights and the y-coordinates of the points P_3 to P_6 are correlated 100 %. By a spatial transformation with five degrees of freedom (3 shifts, 1 rotation, 1 scale) using the given coordinates of the two gable points the other points can be determined together with the 18 x 18 covariance matrix of the 3D-model. The 5 parameters can uniquely be derived taking the identity of the heights for the gable points into account. The projection of these 6 points into the image with the 6 orientation parameters (X_o, Y_o, Z_o, Ω, Φ, K) of the camera yields 12 image coordinates, together with their covariance matrix. Here only Z_o, Ω ,Φ and K need to be treated as random variables with appropriate standard deviations (e. g. $\sigma_{Zo} = 0.1\ Z_o$, $\sigma_\Omega = \sigma_\Phi = \sigma_K = 3\ °$, which is a pessimistic assumption). The resulting 12 x 12 covariance matrix i. g. has rank 6. Observe that this covariance matrix actually represents the non-rigidity of the model and the uncertainty of the approximate values for the orientation parameters.

4.2 The uncertainty of the straight edges

The uncertainty of the straight edges extracted from the image can be derived from the extraction process itself. The final result of straight edge segment, as described in sect. 3, actually results from fitting a straight line through a set of edgels taking the individual weights $w_i = s$ into account. In a local coordinate system (u,v) (cf. Fig. 5), which lies near the principle axes of the edgels, the covariance matrix for the parameters a and m of the straight line $v = a + m\ u$ is diagonal and given by

$$\hat{\underline{C}}(\underline{a},\underline{m}) = \hat{\underline{\sigma}}_o^2\ \mathrm{Diag}(\ 1/\Sigma w_i\ ,\ 1/\Sigma u_i^2 \cdot w_i\ )$$

where $\underline{\sigma}_o^2$ is an estimate for variance factor

$$\hat{\underline{\sigma}}_o^2 = \Sigma\ w_i\ (\underline{v}_i - \hat{\underline{a}} - \hat{\underline{m}}\ u_i)^2\ /\ (n - 2) = \underline{\mu}_2\ /\ (n-2)$$

where the sums are taken for all n edgels. The uncertainty of the u-coordinate in this calcuation can be neglected. The v-coordinates of the end points will be correlated due to the common factor $\hat{\underline{m}}$. The u-coordinates have an accuracy which can be explained merely by rounding errors, thus can be treated as uncorrelated with standard deviation $\sigma_u = 1/\sqrt{12}$. This assumption will be modified in sect. 5. Thus the covariance matrix and the weight matrix of the coordinates of the starting and the end points of the edges in the individual (u,v) coordinate system have the following structure:

$$C = D\begin{bmatrix}\underline{u}_s\\ \underline{v}_s\\ \underline{u}_e\\ \underline{v}_e\end{bmatrix} = \begin{bmatrix}\sigma_{us}^2 & 0 & 0 & 0\\ 0 & \sigma_{vs}^2 & 0 & \sigma_{vsve}\\ 0 & 0 & \sigma_{ue}^2 & 0\\ 0 & \sigma_{vevs} & 0 & \sigma_{ve}^2\end{bmatrix},\quad W_{(u,v)} = \begin{bmatrix}W_{us} & 0 & 0 & 0\\ 0 & W_{vs} & 0 & W_{vsve}\\ 0 & 0 & W_{ue} & 0\\ 0 & W_{vevs} & 0 & W_{ve}\end{bmatrix}$$

Transforming the edge back into the image coordinate system yields the covariance and weight matrices

$$D(\underline{r}_s,\ \underline{c}_s,\ \underline{r}_e,\ \underline{c}_e) = R \cdot C \cdot R^T\ ,\quad W_{(r,c)} = R \cdot W_{(u,v)} \cdot R^T$$

with the rotation matrix R depending on the direction Φ of the individual edge

$$R = \begin{bmatrix} R_\Phi & 0 \\ 0 & R_\Phi \end{bmatrix}, \text{ with } \quad R_\Phi = \begin{bmatrix} \cos\Phi & \sin\Phi \\ -\sin\Phi & \cos\Phi \end{bmatrix}.$$

5. The Matching Criteria

The ultimate goal of our task is to determine the position of the two control points or just one of them. A reasonable requirement is this location to be reliable in the sense that inaccuracies of model or image edges on one hand and false matches on the other hand do not deteriorate the coordinates of the control points too much. This notion of reliability has been developed by BAARDA (1967, 1968) for the use in geodetic networks, a review of the theory is given by FÖRSTNER (1987). We will use it here as a check whether the search for further matches between model and image edges can be terminated or not. As this test in principle has to be performed at all nodes in the search tree, we will elaborate on it separately from the discussion of the search strategy.

Assume a certain list of matched edges $\{(a_k, b_k)\}$ is hypothesized to be acceptable then we can determine the optimal transformation T with parameters y using the nonlinear model

$$E[T(\underline{a}_k; p)] = E[\underline{b}_k]$$

or after linearization

$$\underline{x} + \underline{v} = A\,\hat{\underline{y}}, \qquad C_{xx}$$

with

$\underline{x} = (\underline{x}_k)^T$ the vector of the observations

$\underline{x}_k = T(\underline{a}_k; p^{(o)}) - \underline{b}_k$
the differences between image edge and predicted model edge

$p^{(o)}$ approx. values for the transformation parameters

$\underline{a}_k, \underline{b}_k$ the 4 coord. for the model and the image edge resp.

$\underline{v}_k$ the corrections to the observation $\underline{x}_k$

A the partial derivatives of T with respect to the unknown parameters y

$\hat{\underline{y}}$ the unknown corrections of the parameters p

C_{xx} the covariance matrix of the observations $\underline{x}_k$

The transformation in our case simply can be assumed to be a shift in row- and column-direction. Then the approximate values for the other orientation parameters influence the tests. Thus alternatively, their influence could be eliminated by introducing them in the transformtation T, e.g. in case one is not sure whether the assumed standard deviations are chosen large enough or if one expects the deviations to be too large.

The covariance matrix C_{xx} consists of two parts: $C_{xx} = C_{aa} + C_{bb}$. C_{aa} represents the covariance matrix of the transformed model edges which can be derived by error propagation after linearization of T.

C_{bb} needs some elaboration. It represents the precision of the image edge $\underline{b}_k$, under the assumption that the model holds. Here one has to consider the fact, that model edges not necessarily completely show in the image. Actually only four cases are of interest which are shown in Fig. 6.

Obviously a test on the identity of starting and end points based on parameters derived so far has to be performed first in order to classify the type of match. This test can be split into two phases and refers to the local (u,v) system of the edge:

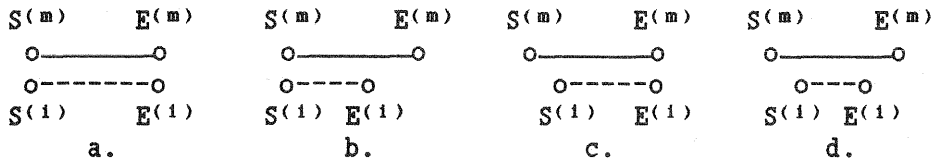


Fig. 6 Possible matches between model edge (o——o) and image edge (o---o)
a. $S^{(m)} = S^{(i)}$, $E^{(m)} = E^{(i)}$
b. $S^{(m)} = S^{(i)}$, $E^{(i)}$ between $S^{(m)}$ and $E^{(m)}$
c. $E^{(m)} = E^{(i)}$, $S^{(i)}$ between $S^{(m)}$ and $E^{(m)}$
d. $S^{(i)}$ and $E^{(i)}$ between $S^{(m)}$ and $E^{(m)}$

a. test of v-coordinates (across the edge). If not both v-coordinates of the image edge are close enough to the predicted model edge, the match is rejected.
b. test of u-coordinates (along the edge).
   - test on the identity of the u-coordinates for both points separately
   - classification of the link at starting and end point according to Fig. 6. Thus if the image edge significantly juts out over the predicted model edge at at least one side the match is rejected.

In case one of the two end points of an accepted image edge is proven to lie between the endpoints of the predicted model edge the standard deviation for the u-coordinate of that point is set to a high value (e. g. 10^3) or the weight is set to 0. This modification of the covariance matrix or the weight matrix reflects the fact the point in concern only being linked to the model edge with its v-coordinate and leaves the structure of the actual estimation procedure unchanged.

Now the checking at each node can be described in the following way:

1. Generate a valid partial solution based on heuristic information (cf. sect. 6). This consists of a list of already accepted matches $\{k\} = \{(\underline{a}_k, \underline{b}_k)\}$, which is $\{0\}$ in the beginning, and a new match $l = (\underline{a}_l, \underline{b}_l)$.
2. Classify the new match l according to the discussion on Fig. 6. If the match is rejected goto 1, else continue.
3. Check the validity of the global match including l using the total fit of the model and the image edges. If this test fails eliminate the reason and goto 1, else accept the match l and continue.
4. Determine the reliability of the control points with respect to mismatches. If the reliability is not sufficient goto 1 else stop.

Remarks:

a. The check in step 3 can be omitted if one can rely on the checks of step 2, which usually is the case.
b. The rejection of the test in step 3 may result from wrong matches which however have passed the tests earlier. Thus a rejection of this test may lead to a test on (possibly all) earlier matches with the possibility to reject one or several of them.
c. The reliability of the result can be based on the theoretical sensitivity ($\delta_o \cdot \sigma_y$), which is known to be an upper bound of the effect of non detectable errors onto the estimate y. It can be based on two different sets of alternative hypothesis:
   1. Complete matches are assumed to be wrong, thus all coordinates of one match are treated as a "joint observation" (SARJAKOSKI 1986). This is a straight forward line of thought. Here the sensitivity measure answers the question: To what extent (in pixels) can the resultant coordinates change if an undetected matching error in one of the matches occured? This is an intuitive "what-happens-if"-type of test. The theoretical sensitivity factor can be derived from the formulas of

the multivariate version of the reliability theory (cf. FÖRSTNER 1983, 1989).

2. If only individual coordinates are assumed to be contaminated by errors, thus the matches are assumed to be correct, the numerical procedure is simplified. How far the actual values differ from case 1 needs special investigations.

6. The Matching Procedure

The matching procedure consists of a backtracking tree search (cf. SHAPIRO/HARALICK 1987) or an exploration of an interpretation tree (cf. GRIMSON/LOZANO-PEREZ 1987) which is shown in Fig. 7.

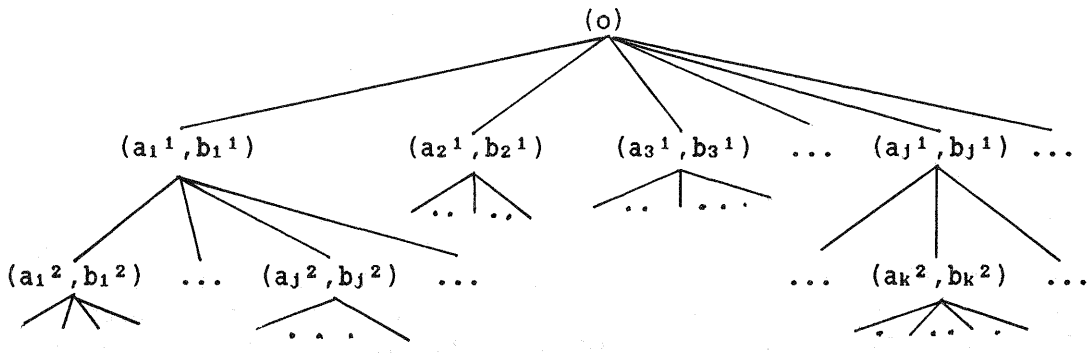


Fig. 7 Interpretation tree

Each node in level 1 represents a possible match (a_m^1,b_m^1) between a model edge a_m^1 and an image edge b_m^1. The path from the root node to a node in level n in the tree represents the above mentioned list of hypothesized matches $\{k\} = \{ak,bk)\}$ with $k = (m^1)$ and $1 = 1,...,n$. The complexity of the search is exponential if no heuristics is used, specifically if we have m model edges to which i image edges may correspond the tree has m^i nodes.

The search can heavily be reduced if not all nodes have to be explored. This can be achieved by sorting the model and the image edges according to their importance for the determination of the cooordinates of the control points and by exploring the tree in a depth first manner, which then corresponds to a best first strategy for finding the solution. As sorting criterium for the model edges again the theoretical sensitivity with respect to potential mismatches can be used, as model edges with a potentially high influence onto the result should be treated first. The sorting criterium for the image edges can be based on the expected gain in precision and thus could be reduced to the closeness of the image edge to the control points. Another means for reducing the search is to clip whole subtrees by putting an upper bound on the global fit between model and image edges. Other possibilities to increase search efficiency can be found but have to tested by simulation studies.

Altogether the presented concept forms a basis for a closed procedure for model based location of objects. It also could be used for matching two edge based image descriptions, e. g. for registration of satellite imagery. Though it has been motivated by locating 3D objects in an image it also could be used for locating 3D objects in space, then requiring special procedures to determine approximate values for the orientation parameters. On the other hand investigations are necessary to test the efficiency of the statistical approach and to check in how far approximations may lead to satisfying results.

References:
IEEE T-PAMI IEEE Transactions on Pattern Analysis and Machine Intelligence

BAARDA W. (1967): Statistical Concepts in Geodesy, Publications in Geodesy, vol. 2. no. 4, Netherlands Geodetic Commission, Delft, 1967

BAARDA W. (1968): A Testing Procedure for Use in Geodetic Networks, Publications in Geodesy, vol. 2. no. 5, Netherlands Geodetic Commission, Delft, 1968

BRADY J. M. (1981): Computer Vision, North-Holland, 1981

BURNS J. B., HANSON A. R. and RISEMAN E. M. (1986): Extracting Straight Lines, IEEE T-PAMI, vol. 8, no. 4, 1986, pp. 425-455

CANNY J. (1986): A Comutational Approach to Edge Detection, IEEE T-PAMI, vol. 8, no. 6, 1986, pp. 679-698

FAUGERAS O. D. and HEBERT M. (1987): The Representation, Recognition, and Positioning of 3-D Shapes from Range Data, in KANADE(1987), pp. 301-353

FÖRSTNER W. (1983): Reliability and Discernability of Extended Gauß-Markov-Models, Deutsche Geod. Komm., A 98, pp. 79-103

FÖRSTNER W. (1986): A Feature Based Correspondence Algorithm for Image Matching. International Archives of Photogrammetry and Remote Sensing Vol. 26-3/3, Rovaniemi 1986.

FÖRSTNER W. (1987): Reliability Analysis of Parameter Estimation in Linear Models with Applications to Mensuration Problems in Computer Vision, Computer Vision, Graphics and Image Processing 40, 1987, pp. 273-310

FÖRSTNER W. (1988): Statistische Verfahren für die automatische Bildanalyse und ihre Bewertung bei der Objekterkennung und -vermessung, Habilitationsschrift, Stuttgart 1988

FÖRSTNER W. (1989): Use of Uncertain Geometric Relationships for Object Location in Digital Images, to appear in Advances in Spatial Reasoning (Ed. S. CHEN), ALBEX 1989

FÖRSTNER, W., GÜLCH, E. (1987): A Fast Operator for Detection and Precise Location of Distinct Points, Corners and Centres of Circular Features. Proc. of Intercommission Conference on Fast Processing of Photogrammetric Data, Interlaken 1987.

FUA P. and HANSON A. J. (1987): Resegmentation Using Generic Shape: Locating Cultural Objects, Pattern recognition Letters 5, 1987, pp. 243-252

GRIMSON W. E. L. and LOZANO-PEREZ T. (1984): Model-Based Recognition and Localization from Sparse Range or Tactile Data, Int. J. Robotics res., vol. 3, no. 3, 1984, pp. 3-34

GRIMSON W. E. L. and LOZANO-PEREZ T. (1987): Localizing Overlapping Parts by Searching the Interpretation Tree, IEEE T-PAMI, Vol. 9, No. 4, 1987, pp. 469-482

HORAUD R. (1987): New Methods for Matching 3-D Objects with Single Perspective Views, IEEE T-PAMI, vol 9, no. 3, 1987, pp. 401-412

KANADE T. (1987, Ed.): Three-Dimensional Machine Vision, Kluwer Academic Publishers, 1987

NEVATIA R. and BABU K. R. (1980): Linear Feature Extraction and Description, Comp. Graph. and Image Proc., vol. 13, 1980, pp. 257-269

SARJAKOSKI T. (1986): Use of Multivariate Statistics and Artificial Intelligence for Blunder Detection, ACSM-ASPRS Annual Convention, Techn. Papers, Volume 4, Washington D. C. 1986, pp. 265-274

SHAPIRO L. G. and HARALICK R. M. (1987): Relational Matching, Applied Optics, vol. 26, 1987, pp. 1845-1851

SMITH R., SELF M. and CHEESEMAN P. (1987a): Estimating Uncertain Spatial Relationships in Robotics, Uncertainty in Artificial Intelligence, vol. 2, 1987, North-Holland

SMITH R., SELF M. and CHEESEMAN P. (1987a): A Stochastic map for Uncertain Spatial Relationships, Proc. of the 4th Int. Symp. on Robotics Research, 1987, MIT Press