In: Burkhardt, K.H.; Höhne, B.; Neumann, B.; (Hrsg.), Mustererkennung 1989, 11. DAGM-Symposium, Hamburg, Band 219 der Reihe Informatik-Fachberichte, S. 457-464. Springer, 1989.

Object Location Based on Uncertain Models

Monika Sester and Wolfgang Förstner
Institut für Photogrammetrie – Universität Stuttgart
Keplerstraße 11, D-7000 Stuttgart 1

Abstract

The paper describes a concept for object location, when not only image features but also the model description is uncertain. It contains a method for probabilistic clustering, robust estimation and a measure for evaluating both, inaccurate and missing image features. The location of topographic control points in digitized aerial images demonstrates the feasibility of the procedure and the usefulness of the evaluation criteria.

1 Introduction

Object location is a central issue in Computer Vision. The task is to determine the pose, i.e. the position and the orientation of an object with respect to a reference frame, for which a model is known either derived from a CAD-system, (cf. GRIMSON/LOZANO-PEREZ 1984,1987, FAUGERAS/HEBERT 1987, HORAUD 1987), sensed from a prototype (FAN 1988), or, more challenging, described by a set of rules, which form a generic model of the object. In this case not only the pose i. e. few parameters are unknown, but also the individual structure of the object (FUA/HANSON 1987).

Main research issues are the computational complexity of the matching problem, which requires strong constraints or heuristics to lead to practially acceptable solutions and the problem of representing 3-dimensional shapes which are suitable for being derived from both CAD-systems and digital images automatically under broad conditions.

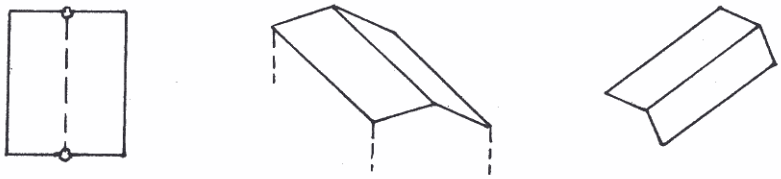
This paper deals with the problem of *uncertainty* encountered during object location from digital images. Uncertainty of the raw data, of the feature extraction, about the used thresholds, about the assumed model, due to wrong or missing correspondencies between model and image features altogether result in uncertain values for the pose parameters. It is the aim of the paper to show, that for this special task of object location the tools provided by mathematical statistics and probability theory are sufficient to describe the uncertainty of the pose parameters in a compact form. The quality of the pose estimation can be used as a means for *selfdiagnosis*, a prerequisite for any automated system to be used in practice.

In contrast to most other approaches, also the uncertainty of the model is taken into account. Moreover, the theory enables to test the validity of the model and to evaluate the effect of both imprecise and missing image features onto the result.

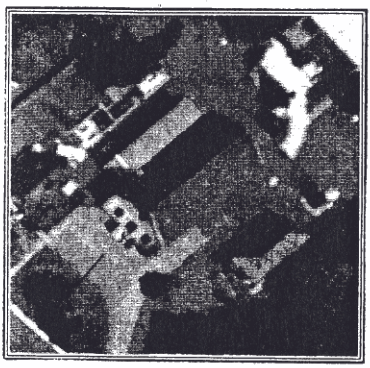
The motivation for this study resulted from a task at the photogrammetric department of the *Landesvermessungsamt Nordrhein-Westfalen*. For their orthophotoproduction they have built up a data base of more than 20000 topographic control points, mainly gable points of house roofs. They use their location in the image for determining the orientation of the aerial photos (scale appr. 1:12000). The X,Y and Z world coordinates of the gable points and a sketch of the roof in orthogonal projection are given. A first prototype program to locate the control points automatically reveiled the difficulty to transparently set thresholds in the image analysis and the matching procedure which required a theoretically more founded setup described in this paper.

The whole identification procedure consists of 4 steps (cf. Fig. 1).

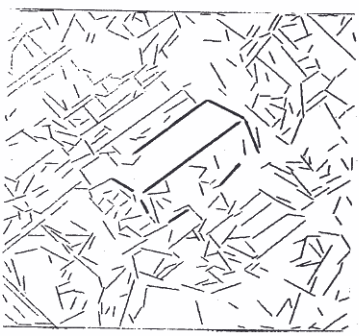
1. *Interpretation* of the 2-D-sketches, resulting in a 3-D-description. It not only contains the coordinates in the world coordinate system, but also their uncertainty as the sketches are not fully in scale (cf. Fig. 1a → 1b).
2. *Projection* of the 3-D-model into the aerial image using the appropriate values for position and orientation of the camera from a flight plan leading to a 2-D-wire frame model in the image, again containing the uncertainty, now of both the sketch and the appropriate orientation parameters (cf. Fig. 1b → 1c).



a. given sketch, orthogonal projection b. interpreted sketch, 3D-model c. projected model, 2D-wire model



d. scanned image section, 240 x 240 pixels à 20 μ



e. extracted edges (thin), matched edges (thick)

Figure 1: Example for matching roof model to image edges

3. *Extraction* of straight line segments from a digitized subsection of the areal image, leading to a list of line segments, also here containing information about the geometric uncertainty due to the image analysis (cf. Fig. 1d → 1e).
4. *Matching* the image features to the model features taking the geometric relations and the uncertainty into account (cf. Fig. 1c ⇔ 1e).

Section 2 outlines the theoretical background of the approach. Sections 3 and 4 describe the chosen representation of model and image features. Section 5 contains the used matching procedure. An example of the procedure in section 6 demonstrates the feasibility of the approach and the usefulness of the chosen evaluation criteria.

2 Conceptual Background

The ultimate goal of our task is to determine the position of the control points, i. e. the gable points. A reasonable requirement is the location to be reliable in the sense that inaccuracies of model or image features on one hand and false matches on the other hand do not deteriorate the coordinates of the controlpoints too much. This notion of reliability has been developed by BAARDA 1967,1968 for the use in geodetic networks and seems to be appropriate here too, a review of the theory is given by FÖRSTNER 1987.

2.1 The Functional and the Stochastical Model

Assume a certain list of matched features $\mathbf{a}_o = \{a_k\}$ is hypothesized to be acceptable then we can determine the optimal transformation T with parameters $\mathbf{p}$ using the nonlinear model (stochastical variables are underlined)

$$E(\underline{\mathbf{a}}_o) = T(\mathbf{p}) = T(\mathbf{p}_t, \mathbf{p}_f, \mathbf{p}_l) \quad (1)$$

$$E(\underline{\mathbf{p}}_f) = \mathbf{p}_f \quad (2)$$

$$E(\underline{\mathbf{p}}_l) = \mathbf{p}_l \quad (3)$$

or after linearisation

$$\underline{\mathbf{x}} + \underline{\mathbf{r}} = \mathbf{A}\hat{\underline{\mathbf{y}}} \quad or \quad \begin{pmatrix} \underline{\mathbf{x}}_o \\ \underline{\mathbf{x}}_f \\ \underline{\mathbf{x}}_l \end{pmatrix} + \begin{pmatrix} \underline{\mathbf{r}}_o \\ \underline{\mathbf{r}}_f \\ \underline{\mathbf{r}}_l \end{pmatrix} = \begin{pmatrix} \mathbf{A}_t & \mathbf{A}_f & \mathbf{A}_l \\ 0 & \mathbf{I} & 0 \\ 0 & 0 & \mathbf{I} \end{pmatrix} \begin{pmatrix} \hat{\underline{\mathbf{y}}}_t \\ \hat{\underline{\mathbf{y}}}_f \\ \hat{\underline{\mathbf{y}}}_l \end{pmatrix} \tag{4}$$

with

$\underline{\mathbf{x}}_o = (\underline{\mathbf{x}}_{ok}^T)^T$	the vector of observations
$\underline{\mathbf{x}}_{ok} = \underline{\mathbf{a}}_k - T(\mathbf{p}^{(o)})$	the differences between image features and the predicted model features
$\mathbf{p}^{(0)} = (\mathbf{p}_t^{(0)T}, \mathbf{p}_f^{(0)T}, \mathbf{p}_l^{(0)T})^T$	approximate values for the transformation parameters for the translation ($\mathbf{p}_t$), the form ($\mathbf{p}_f$) and the location parameters ($\mathbf{p}_l$)
$\hat{\underline{\mathbf{y}}} = (\hat{\underline{\mathbf{y}}}_t^T, \hat{\underline{\mathbf{y}}}_f^T, \hat{\underline{\mathbf{y}}}_l^T)^T$	the estimated corrections to the approximations
$\mathbf{a}_k$	the parameters describing the image feature k
$\underline{\mathbf{r}}$	the corrections to the observations $\underline{\mathbf{x}}$
$\mathbf{A}_o = (\mathbf{A}_t, \mathbf{A}_f, \mathbf{A}_l)$	the partial derivations of T with respect to the unknown parameters

Eqs.1 to 4 contain the geometric relationship between the image features and the unknown parameters, in nonlinear and linearized form. They represent the projection of the object into the image. The unknown parameters describe the form of the object and the orientation of the image with respect to the object; the task is equivalent to locating the object with respect to the image coordinate system. The additional eqs. 2 and 3 contain the a priori knowledge about the form and the orientation parameters. The *functional model* thus is a statement about the expectation about the observed values, or more technically about the first moments of the stochastical variables associated with the uncertain input values.

The *stochastical model* describes the uncertainty of the observed image features $\mathbf{a}_k$ and of the approximate values $\mathbf{p}_f^{(0)}$ and $\mathbf{p}_l^{(0)}$. The uncertainty of the stochastical variables is represented by their second moments or their variances:

$$D(\underline{\mathbf{x}}) = D(\underline{\mathbf{a}}_o, \underline{\mathbf{p}}_f, \underline{\mathbf{p}}_l)^T = diag\,(D(\underline{\mathbf{a}}_o), D(\underline{\mathbf{p}}_f), D(\underline{\mathbf{p}}_l)) = diag\,(\mathbf{C}_{oo}, \mathbf{C}_{ff}, \mathbf{C}_{ll}) = \mathbf{C}_{xx} \tag{5}$$

In case the covariance matrices $\mathbf{C}_{ff}$ and $\mathbf{C}_{ll}$ are diagonal the standard deviations σ_{p_f} and σ_{p_l} fully describe the uncertainty of the model and the approximate values. The matix $\mathbf{C}_{oo}$ in general will not be diagonal, describing the interdependencies between the geometric feature values derived from the image.

The model eq. 4 is general enough to represent a large class of object location tasks. We will specify the individual parts later.

2.2 Precision and Sensitivity of Estimated Parameters

The parameters can now be estimated from a best linear unbiased estimation procedure:

$$\hat{\mathbf{y}} = (\mathbf{A}^T\mathbf{C}_{xx}^{-1}\mathbf{A})^{-1}\,\mathbf{A}^T\mathbf{C}_{xx}^{-1}\,\mathbf{x} \tag{6}$$

The influence of random errors onto the result, i.e. the *precision* can easily be derived by error propagation. The covariance matrix

$$\mathbf{C}_{yy} = (\mathbf{A}^T\mathbf{C}_{xx}^{-1}\mathbf{A})^{-1} = \begin{pmatrix} \mathbf{C}_{tt} & \mathbf{C}_{tf} & \mathbf{C}_{tl} \\ \mathbf{C}_{ft} & \mathbf{C}_{ff} & \mathbf{C}_{fl} \\ \mathbf{C}_{lt} & \mathbf{C}_{lf} & \mathbf{C}_{ll} \end{pmatrix} \tag{7}$$

contains the variances of all unknown parameters. This specially holds for the translation parameters $\hat{\mathbf{p}}_t$ and the form parameters $\hat{\mathbf{p}}_f$.

A test on the residuals $\mathbf{r}_o$ reveals outliers in the image features, which may result from a wrong correspondence. A test on the residuals $\mathbf{r}_f$ may reveal errors in the assumed form of the object. For these tests an assumption about the distribution of the stochastical variables is required. A Gaussian distribution seems to be sufficient in all cases.

Even if the derived standard deviations are acceptable the result may be uncertain due to two reasons:

1. The ommittance of one of the image features say $\mathbf{x}_k$ may lead to a large change of the pose parameters

$\hat{\mathbf{y}}_p = (\hat{\mathbf{y}}_t, \hat{\mathbf{y}}_l)$. An upper bound for the influence (indicated by ∇) of image feature $\mathbf{x}_k$ onto an arbitrary function $f(\hat{\mathbf{y}}_p)$ of $\hat{\mathbf{y}}_p$ is given by

$$\nabla_k f(\hat{\mathbf{y}}_p) \leq \sigma_{f(y_p)}\, \mu_k\, t_k = \sigma_{f(y_p)}\, \bar{\delta}_k \tag{8}$$

where

- $\sigma_{f(\mathbf{y}_p)}$ is the standard deviation of $f(\hat{\mathbf{y}}_p)$
- μ_k describes the geometric sensitivity of the result with respect to the k-th feature and - for simplicity at the moment assuming no form parameters to be estimated - can be derived from

$$\mu_k = \lambda_{max}[\mathbf{C}_{\hat{x}_k\hat{x}_k}\,(\mathbf{I} - \mathbf{C}_{\hat{x}_k\hat{x}_k})^{-1}] \tag{9}$$

  with $\mathbf{C}_{\hat{x}_k\hat{x}_k} = \mathbf{A}_k\mathbf{C}_{yy}\mathbf{A}_k^T$, where $\mathbf{A}_k$ is the k-th submatrix of $\mathbf{A}_o = (\mathbf{A}_t, \mathbf{A}_f, \mathbf{A}_l) = (\mathbf{A}_{o1}^T, \ldots, \mathbf{A}_{ok}^T, \ldots)^T$ (cf. FÖRSTNER 1983)
- t_k is the statistic for testing the significance of the residual $\mathbf{r}_k$ with respect to wrong correspondencies (namely the squareroot of a Fisher-teststatistic)
- $\bar{\delta}_k$ is an factor depending on μ_k and t_k, thus on the k-th feature only

Hence, even if σ_f and t_k are acceptable due to a weak geometry the k-th feature may heavily influence the result if μ_k is large. On the other hand even if σ_f and t_k are not acceptable the influence of $\mathbf{a}_k$ onto the result may be small, indicating that the feature may be superfluous. The upper bound $\sigma_f\, \mu_k\, t_k$ thus measures the *empirical sensitivity* of the result. The normalized and dimensionless *empirical sensitivity factor* $\bar{\delta}_k = \mu_k\, t_k$ therefore should be in the range up to 1 – 2.

2. The *theoretical sensitivity* does not depend on the data and measures the ability of the result to absorb wrong matches which are undetectable by a test on the residuals $\mathbf{r}_k$. The ability to detect (gross) errors depends not only on the geometry, thus on μ_k, but also on the used test, its degrees of freedom dof, the significance level $S = 1 - \alpha$ and the required minimum power β_0 for the test. We obtain the upper bound for the influence of an undetectable error to be

$$\nabla_{0k} f(\mathbf{y}_p) \leq \sigma_{f(y_p)}\, \mu_k\, \delta_0 = \sigma_{f(y_p)}\, \bar{\delta}_{0k} \tag{10}$$

with $\delta_0 = \delta_0(\alpha_0, \beta_0, dof)$, e. g. being $\delta_0 = 4.13$ for $dof = 1, \alpha = 0.001$ and $\beta_0 = 0.80$ (cf. BAARDA 1967, 1968, FÖRSTNER 1987).

The relevance of this measure is its ability to indicate *missing* image features, which are necessary for the determination of the pose parameters, as then the *theoretical sensitivity factors* $\bar{\delta}_{0k}$ of the *other* features may easily reach values greater than 4 – 5. The relevance of the standard deviation of the pose estimates, measuring the influence of the random errors only, is reduced in this case. A similar reasoning can be applied to the approximate values specifying form and orientation. This leads to statements about the sensitivity of the result with respect to gross errors in the model or in the assumed orientation of the camera.

The upper bounds eqs. 8 and 10 contain information on the effect of random and gross errors, as well as missing data. As they are in the dimension of the resultant parameters, they are easily interpretable.

3 Representation of Uncertain Models

The basic idea of representing uncertain models is to describe the object in a parametric form and to specify the uncertainty of the parameters by their standard deviation. Though we only discuss the models used in a specific application the potential of the method will become clear (cf. SMITH ET AL. 1987).

In our case the model shows up in three different versions, which can be derived by geometric transformations each involving a well defined set of parameters:

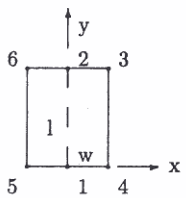
a: The planar coordinates of the nodes in the sketch can be expressed by a few parameters, e.g. the width of the roof (cf. Fig. 2). The scale can directly be taken into account due to the given, and fixed coordinates of the two gable points, which therefore serve as an intuitive coordinate system. The imprecision of the sketches leads to reasonable standard deviations of appr. 10–15% of the lengths. The Z-values of the house depend on the slope of the roof which is unknown. Due to this lack of knowledge we allow each roof-plane to be of type horizontal, positive slope or negative slope unless geometric conditions form constraints on

the choice. The interpretation of the sketches, which is necessary here, will not be discussed. We assume that one choice, among several has been made. Then the slopes are additional free parameters specifying the form. Their expected value is assumed to be approx. $\pm 45^\circ$ or 0°, the standard deviation to be $\sigma = 15^\circ$.

The 3D-coordinates (X_i, Y_i, Z_i) of the model in full scale then depend on the parameters $\mathbf{p}_f$:

$$\begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \\ \mathbf{Z} \end{pmatrix}_i = B_i(\mathbf{p}_f) = \begin{pmatrix} \mathbf{B}_{xi} \\ \mathbf{B}_{yi} \\ \mathbf{B}_{zi} \end{pmatrix} \mathbf{p}_f \tag{11}$$

For the simple roof in Fig. 2 we obtain with $\mathbf{X} = (X_i), \mathbf{Y} = (Y_i)$ and $\mathbf{Z} = (Z_i)$ and the two parameters



i	B_{xi0}	B_{xi1}	B_{xi2}	B_{yi0}	B_{yi1}	B_{yi2}	B_{zi0}	B_{zi1}	B_{zi2}
1	0	0	0	0	0	0	0	0	0
2	0	0	0	1	0	0	0	0	0
3	0	-1	0	0	0	0	0	$-s$	$-w$
4	0	-1	0	1	0	0	0	$-s$	$-w$
5	0	+1	0	0	0	0	0	$-s$	$-w$
6	0	+1	0	1	0	0	0	$-s$	$-w$

Figure 2: Example for a parametric description of a sketch; w = width of the roof, l = length of the roof

$p_{f_1} = width = w$ and $p_{f_2} = slope$ and using $p_{f_0} = l$ the elements in the table of Fig. 2. $\mathbf{B}$ and $\mathbf{C}_{ff}$ symbolically represent the 3D-model.

b: The transformation T_l^w of this 3-D-model into the world coordinate system requires no new uncertain parameters.

c: The projection T_w^m of the 3-D-world model into the image depends on 6 parameters from which the translation $p_t = (X_0, Y_0)$ is assumed to be free (or with very large standard deviations, measured in pixels). The four other parameters p_l, namely scale and orientation, are assumed to have a precision of 5–10 % and $2^0 - 4^0$ resp. Thus the image coordinates of the projected model can be derived for each point i by

$$\mathbf{b}_i = \begin{pmatrix} x \\ y \end{pmatrix}_i = T_w^m(\mathbf{p}_t, \mathbf{p}_l)\, T_l^w\, B_i(\mathbf{p}_f) \tag{12}$$

For a point pair $\mathbf{a}_k = (\mathbf{b}_i, \mathbf{b}_j)$, forming an edge in the projected model this after linearization yields the right hand side of the first line in eq. 4.

Without additional knowledge the parameters then would be free. The a priori knowledge of the form and the orientation therefore is expressed by eqs. 2, 3 or the 2nd and 3rd line in eq. 4.

4 Representation of Straight Edges Extracted from Images

Any extraction scheme can be used for deriving straight edge elements (cf. e. g. BURNS ET AL. 1986, NEVATIA/BABU 1980). The uncertainty of the straight edges can usually be derived from the extraction process itself. We use a weighted least squares fit through the extracted edge elements, determined to subpixel position. The weights w_i are proportional to the squared value of the gradient. In a local coordinate system (u,v) (cf. Fig. 3), which lies near the principle axes of the edge elements the covariance matrix for

Figure 3: Representation of edge and local coordinate system (u,v)

the parameters a and m for the straight line $v = a + mu$ is diagonal and given by:

$$D\begin{pmatrix} \hat{a} \\ \hat{m} \end{pmatrix} = \sigma_0^2 \begin{pmatrix} 1/\sum w_i & 0 \\ 0 & 1/\sum u_i^2 w_i \end{pmatrix} \tag{13}$$

where the sums are taken over all edge elements. The uncertainty of the u-coordinates in this calculation can be neglected. The v-coordinates of the end points will be correlated due to the common factor $\hat{m}$. The u-coordinates of the edges have an accuracy which can be explained merely by rounding errors, thus can be treated as uncorrelated with standard deviation $\sigma_u = 1/\sqrt{12}$ pixels.

The extracted edge segment now is represented by three groups of observations:

1. The pair of v-coordinates with:

$$\mathbf{C}_{vv} = D\begin{pmatrix} \underline{v}_s \\ \underline{v}_e \end{pmatrix} = \begin{pmatrix} \sigma_{v_s}^2 & \sigma_{v_s v_e} \\ \sigma_{v_s v_e} & \sigma_{v_e}^2 \end{pmatrix} \quad or \quad \mathbf{W}_{vv} = \mathbf{W}\begin{pmatrix} \underline{v}_s \\ \underline{v}_e \end{pmatrix} = \begin{pmatrix} w_{v_s}^2 & w_{v_s v_e} \\ w_{v_s v_e} & w_{v_e}^2 \end{pmatrix} = \mathbf{C}_{vv}^{-1} \tag{14}$$

2./3. The u-coordinate system of starting point and end point resp.:

$$D(\underline{u}_s) = \sigma_{u_s}^2 \quad \text{or} \quad w_{u_s} = 1/\sigma_{u_s}^2 \qquad \text{and} \qquad D(\underline{u}_e) = \sigma_{u_e}^2 \quad \text{or} \quad w_{u_e} = 1/\sigma_{u_e}^2 \tag{15}$$

The weights will be modified in the robust estimation, allowing partial matches of image and model segments (4 cases). The complete weight matrix $\mathbf{W}^{(u,v)}$ of $(\underline{u}_s, \underline{v}_s, \underline{u}_e, \underline{v}_e)^T$ is transformed into the xy-coordinate system using the individual rotation matrix $\mathbf{R} = \begin{pmatrix} \mathbf{R}_\phi & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_\phi \end{pmatrix}$ with $\mathbf{R}_\phi = \begin{pmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}$

$$\mathbf{W}^{(x,y)} = \mathbf{R}\mathbf{W}^{(u,v)}\mathbf{R}^T \tag{16}$$

5 The Matching Procedure

The matching of the image edge segments to the model edges is performed in a two step procedure:

- The approximate position of the projected model in the image is determined via a *probabilistic clustering* technique, leading to a small set of candidate matches.
- The final correspondence is established via a *robust estimation*, which corresponds to the well known relaxation techniques.

5.1 Probabilistic Clustering for finding Candidates for Correspondencies

The approximate position of a reference point of the projected model is determined via clustering. In contrast to the usual procedure for clustering the varying uncertainty of the hypothesized position is taken into account when filling the accumulator. The final cluster is a discretized version of the probability density function (p.d.f) $p_{xy}(x, y)$ of the reference point.

For each match of image feature i with model feature j we get a range of expected positions for the reference point, as the image segments in general are shorter than the model segments. For a fixed position t, $t \in [0, 1]$ the p.d.f. is given by

$$p_{xy}(x, y|i, j, t) = \frac{1}{2\,\pi\,\sigma_{ijt}^2} \exp\left\{-\frac{1}{2}\frac{(x - x(t))^2 + (y - y(t))^2}{\sigma_{ijt}^2}\right\} \tag{17}$$

where the variance $\sigma_{ijt}^2 = (1/2 + (s_{ijt}/l_i)^2)\sigma_i^2 + \sigma_{s_{ijt}}^2$ depends on the distance s_{ijt} of the midpoint of the image segment from the reference point, its variance $\sigma_{s_{ijt}}^2$, the length l_i of the image segment and the uncertainty of its endpoints σ_i^2, neglecting the correlations in this step.

$P_0 = (x(0), y(0))$ and $P_1(x(1), y(1))$ are the extreme positions the reference point may occupy (cf. Fig. 4a). The p.d.f. $p_t(t|i, j)$ is assumed to be an equal distribution $e(t|i, j)$ on the straight line segment between P_0 and P_1. The probability $P_{ij}(i, j)$, that i and j correspond is assumed to be $1/|M|$ where M is the set of all matches. Then the cluster is determined by

$$p_{xy}(x, y) = \sum_{ij \in M} \left[\int_{t=0}^{1} p_{xy}(x, y|i, j, t)\, p_t(t|i, j)\, dt\right] P_{ij}(i, j) = \frac{1}{|M|} \sum_{ij \in M} \int_{t=0}^{1} p_{xy}(x, y|i, j, t)\, e(t|i, j)\, dt \tag{18}$$

$p_{xy}(x, y|i, j, t)$ is approximated by a separable binomial distribution and the integrals by sums over all positions on the grid points of the accumulator between the extrems P_0 and P_1. Examples for the integral in eq. 18 are shown in Fig. 4b (s_{ijt} is approximated by $s_{ij0.5}$ here).

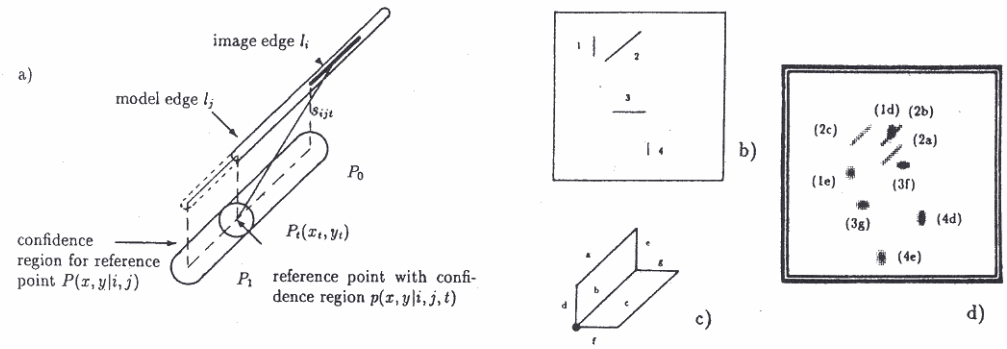


Figure 4: Principle of edge matching (a), example for distribution in accumulator for a few edges (d), selected edges (b), model (c)

With the approximate position of a reference point one easily can derive a set of candidate matches between image and model features. Specifically we obtain a set of straight image segments with a label indicating to which model segment they may correspond. The threshold for this conservative decision has to take the uncertainty of the model into account.

5.2 Robust Estimation and Final Evaluation

The list of preliminary correspondencies now is cleaned using the model eq. 4. As wrong correspondencies may be interpreted as outliers or with large deviations from the mean value, a robust estimation is performed (cf. HUBER, 1981). The principle is simply weighting down those observations which have large residuals or a large test statistic. In detail we use the following iteration scheme

$$\mathbf{W}_{vv}^{(\nu+1)} = \mathbf{W}_{vv}^{(0)}\, f(t^{(\nu)}) \qquad w_{u_s}^{(\nu+1)} = w_{u_s}^{(0)}\, f(t^{(\nu)})\, f(t_{u_s}^{(\nu)}) \qquad w_{u_e}^{(\nu+1)} = w_{u_e}^{(0)}\, f(t^{(\nu)})\, f(t_{u_e}^{(\nu)}) \tag{19}$$

For the first 3-6 iterations and the following 1 or 2 iterations the weight functions

$$f_1(t) = 1/\sqrt{1 + (t/c)^2} \quad \textit{and} \quad f_2(t) = \exp\left\{-0.5(t/c)^2\right\} \tag{20}$$

resp. are used. t, t_{u_s} and t_{u_e} are the square-rooted Fisher-test statistics for the complete segments and the u-coordinates of the starting and the end point resp. c is a critical value, e. g. $c = 3$. The first weighting function leads to a mixture between L_1- and L_2-norm minimization and guarantees global convergency for linear problems, the second cancels the effect of large outliers onto the result.

As starting and end point are treated separately all possible 4 cases of an image segment matching a model segment can be realized. All correspondencies fitting to the final result, i. e. passing the test are assumed to be correct. The final result is then analyzed with respect to precision and sensitivity according to section 2.2. It is accepted if the image features, non detectable errors in the correspondencies and the assumptions contained in the model do have an effect onto the resultant position which does not exceed a prespecified threshold, e. g. a few pixels.

6 Example

From all the image segments in Fig. 1e the clustering process (cf. Fig. 5a) selected 17 segments as candidates (cf. Fig. 5b). The robust estimation selected the 8 segments shown in Fig. 5c. The precision of the reference point was $\sigma_p = 25.9\ \mu m$. The check on the theoretical sensitivity reveals edge segment 7 to have the strongest possible effect onto the result with $\bar{\delta}_{07} = 4.9$ (cf. Table 1), i. e. with a maximum influence of $\nabla_{07} p = \sigma_p \cdot \bar{\delta}_{07} = 126\ \mu m$, which is at the border of being acceptable. In case only the 4 segments in Fig. 5d would have been selected by the robust estimation, the precision of the reference point would

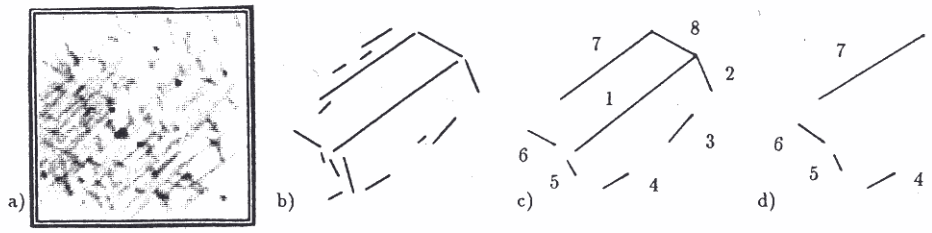


Figure 5: Cluster for approximate position (a), preliminary matches (b), matches selected by robust estimation (c), ficticiously reduced sample (d)

case	precision	sensitivity	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	width	slope
(c)	$\sigma_p = 25.9\mu m$	empirical	0.5	1.1	1.0	1.1	1.6	1.3	1.3	1.0	0.01	0.05
		theoretical	3.1	2.9	2.7	1.4	3.4	3.6	4.9	3.2	0.11	0.79
(d)	$\sigma_p = 81.5\mu m$	empirical	–	–	–	7.0	0.6	0.3	4.0	–	0.8	0.09
		theoretical	–	–	–	14.5	2.9	3.1	14.3	–	1.4	5.0

Table 1: Final evaluation of design for solutions (c) and (d) from Fig. 5. Sensitivity factors $\bar{\delta}_k$ and $\bar{\delta}_{0k}$

have been $\sigma_p = 81.5\ \mu m$. Also, due to the lack of stability errors in the matching may have influenced the result up to $\bar{\delta}_{04} = 14.5$ or up to 1.2 mm, which is appr. 60 pixels. Especially the high value of $\bar{\delta}_{04}$ indicates missing support for the model from the image. It is interesting to note, that also the apriori assumption about the slope of the roof has a significant influence onto the position of the reference point, namely up to 0.4 mm. As the test statistics all were accepted these values at the same time represent upper bounds on the empirical sensitivity of the result.

References

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.

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

FAN, T.-J. (1988) : Describing and Recognizing 3D-objects using Surface Properties, Inst. for Robotics and Intelligent Systems, Report Nr. 237, School of Engineering Univ. of Southern California, L.A., Aug. 1988.

FAUGERAS, O.D., 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. (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.

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

GRIMSON, W.E.L., 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., 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.

HUBER (1981) : Robust Statistics, Wiley, NY, 1981.

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

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

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