# Generic Estimation Procedures for Orientation with Minimum and Redundant Information

*Wolfgang Förstner*

Institut für Photogrammetrie
Universität Bonn
Nußallee 15, D-53115 Bonn 1
e-mail: wf@ipb.uni-bonn.de

## 1 Motivation

Orientation of cameras with minimum and redundant information is the first step in 3D-scene analysis. Compared to image interpretation it looks simple, it seems to be solved in photogrammetry and is expected to be implemented within a few weeks. All experience shows that camera calibration and orientation needs much effort and the solutions provided in photogrammetric textbooks cannot be directly transferred to automatic systems for scene analysis.

The reasons for this situation lie in the hidden complexity of the calibration and orientation tasks.

- *Camera modelling* requires a thorough understanding of the physics of the image formation process and of the statistical tools for developing *and* refining mathematical models used in image analysis. High precision cameras used in aerial photogrammetry have put the burden of obtaining high precision on the manufacturer, leading to the - only partly correct - impression that calibration can be standardized, and thus is simple. The long scientific struggle photogrammetry went through in the 70's, which is not mentioned in today's publications, must now be repeated under much more difficult boundary conditions: non-standardized video cameras, non-standardized applications, the requirement for full automation, therefore the integration of error-prone matching procedures, etc..
- The *3D-geometry of orientation* reveals high algebraic complexity. This is overseen when assuming the calibration and orientation to be known or at least approximately known, as the iterative techniques used in photogrammetry and the spatial intersection (triangulation) - in general - lead to satisfying results. Again, the efforts of photogrammetric research in the 70's and early 80's for generating guidelines for a good design of so-called "photogrammetric blocks", where hundreds and thousands of images are analysed simultaneously for object reconstruction, specifically mapping, has to be invested for the different tasks of 3D-scene reconstruction in computer vision, especially in the area of structure from motion. It is interesting and no accident

that such guidelines are only available for aerial photogrammetric blocks, not for close range applications. The complexity of the 3D-geometry of orientation motivated the numerous publications in the computer vision area on the availability, uniqueness and stability of orientation and reconstruction procedures under various, partly very specific, boundary conditions.

– *Error handling* is a central issue in calibration and orientation of cameras for several reasons.
  - The *correspondence problem* is far from being solved for general cases. Existing solutions have to deal with large percentages of matching errors. This prevents the direct use of classical estimation procedures and makes it necessary to look for robust procedures which, however, make a thorough analysis of the quality of the final result at least difficult, as the underlying theories (!) often only give asymptotic theorems.
  - In case *approximate values* for calibration and orientation are not available or only of poor quality their determination appears to be a far more challenging problem than the refinement via a least squares estimation. The direct solutions, either with minimum or redundant information play a central role, especially in the presence of outliers.
  - *Self-Calibration* is often required where calibration, orientation and generally also scene reconstruction is performed simultaneously, as camera calibration in a laboratory often is not feasible or insufficient. It increases the difficulty of error analysis by at least one order of magnitude as deficiencies in design, modelling and mensuration have to be handled simultaneously and, therefore, generally prevent an algebraic analysis of the system. The difficulty of integrating *all* types of observational values lies in the necessity to formalize the evaluation process in order to adequately handle the different dimensions (pixels, meter, radiants, etc.) of the observations and their influence on the final result.

  Experiences in photogrammetric research give many hints on how to solve the problem of error handling, especially with respect to the quality evaluation based on various statistical tools. Nonetheless, the boundary conditions met in computer vision applications require a new setup of the concepts.

– The final goal of image analysis is *full automation* of all procedures. As calibration and orientation of cameras, due to its well-defined goal, really is much more simple than image interpretation, it seems to be feasible to achieve generic procedures for automatically solving this first step in the analysis chain. Textbooks on photogrammetry, statistical analysis or other related topics, however, often only *present tools not strategies* for solving the problem of parameter estimation, calibration and orientation like many other sub-tasks in image analysis. This is due to the specific engineering expertise which is required to find the appropriate tool combination. This expertise is usually not documented in textbooks, but in internal reports of institutions for training purposes, e. g. for handling complex software packages. Sometimes this knowledge is already formalized in terms of a sequence of rules to be applied.

> Formalization, being a prerequisite for developing generic procedures, is difficult in our context as the various types of errors (cf. subsection 2.2 on error handling) interfere in a nonpredictable manner and no coherent theory is available to justify specific strategies.

This paper is motivated by this deficit in generic and robust procedures for geometric reasoning, calibration and especially orientation. Its aim is to collect the available tools from statistics, specifically for the diagnosis of data and design and for coping with outliers using robust estimation techniques, and to present a generic strategy for data analysis in the context of orientation procedures. The techniques allow an extension towards self-calibration which, however, has to be worked out. The much more difficult problem of designing, i. e. planning mensuration procedure of high robustness, still waits for a solution.

## 2 Problem Statement

Let us assume the model to explicitely describe the observation process

$$E(\boldsymbol{l}) = \boldsymbol{g}(\boldsymbol{\beta}) \tag{1}$$

where the expectation of the $n$ observations $\boldsymbol{l} = \{l_i\}$ via $\boldsymbol{g}$ in general nonlinearily depends on the $u$ unknown parameters $\boldsymbol{\beta} = \{\beta_j\}$. The stochastical properties of the observations are captured by the covariance matrix

$$D(\boldsymbol{l}) = \boldsymbol{\Sigma}_{ll}\,. \tag{2}$$

Should this be the only information available the principle of maximum entropy results in the following full model

$$\boldsymbol{l} \sim N(\boldsymbol{g}(\boldsymbol{\beta}), \boldsymbol{\Sigma}_{ll}) \tag{3}$$

hypothesizing $\boldsymbol{l}$ to be normally distributed. The redundancy of the system is

$$r = n - u\,. \tag{4}$$

The task is to derive estimates $\widehat{\boldsymbol{\beta}}$ from given observational values $\boldsymbol{l}$.

In our context the observations usually are the coordinates of points or the parameters of lines detected and located in the image by an automatic procedure. The relation between corresponding points and/or lines in several images or in object space, also performed automatically, guarantees redundancy in the total process, as several image features generally determine one corresponding object feature.

In case the redundancy equals 0 or in the unlikely case of the observations being consistent, the assumed stochastical properties have no influence on the estimate. The only task then is to invert (1) to obtain $\widehat{\boldsymbol{\beta}} = \boldsymbol{g}^{-1}(\boldsymbol{l}_s)$, where $\boldsymbol{l}_s$ is a subset of $\boldsymbol{l}$ of size $u$.

### 2.1 Error Types

In general, all components of the model will have an influence on the result. The key question is how an automatic system handles errors in these assumptions. One may distinguish three types of errors:

1. *Data errors*, which are errors in the values of $\boldsymbol{l}$, grossly violate assumption (3). They relate to points, lines or other features in the image or in object space where measurements are taken. They may really be mensuration errors, e. g. caused by failures in the detection algorithm or matching errors leading to wrong relations between image and object features. Depending on the complexity of the scene and the quality of the used algorithms the percentage of errors may range between a few and over 80 % of the observed values.
2. *Model errors* refer to all three parts of the model: the functional relationship $\boldsymbol{g}(\beta)$, the covariance matrix $\boldsymbol{\Sigma}_{ll}$ and the type of the distribution, here the normal distribution $N(\cdot\,,\,\cdot)$. Examples for this type of error are manifold:
   - too few, too many or the wrong set of parameters $\boldsymbol{\beta}$, e. g. when using shallow perspective, projectivity or parallel projection;
   - wrong weighting, e. g. when assuming the same accuracy for all detected points;
   - neglected correlations, e. g. in Kalman-filtering; or,
   - wrong assumptions about the distribution, e. g. when handling one-sided errors.

   Observe that data errors and model errors cannot formally be distinguished; as a refinement of the model may always specify the type of error in the observations.
3. *Design or configuration errors* relate to the complete set of functions $\boldsymbol{g} = \{g_i\}$. Such errors cause the estimate $\widehat{\boldsymbol{\beta}}$ to be nonunique in some way. Multiplicity of solutions is the best case of nonuniqueness. Depending on the degree of redundancy we may distinguish at least three cases (cf. the formalization in section 3.2):
   (a) nondeterminable parameters. Critical surfaces of the configuration belong to this class. An example would be a spatial resection with three points and the projection centre sitting on the critical cylinder.
   (b) noncheckable observations or parameters. Here the determination of the parameters may be possible, but errors in the estimated parameters introduced in a Bayesian manner, are not detectable due to a too low redundancy. An example would be a spatial resection with three points in general position.
   (c) nonlocatable errors. Here a test may be able to show discrepancies between the data and the model, but no identification of the error source is possible. An example would be a spatial resection with four points in general position.

We will treat all types of errors in the following; however, concentrate on means for automatically reacting on indications of such errors.

## 2.2 Issues in Error Handling

There are at least three basic questions that automatic procedures need to be able to answer:

1. How sensitive are the results?
   The results may be uncertain due to the large number of errors mentioned above. Evaluating real cases has to cope with the problem that several such errors occur simultaneously. *Instabilities* due to low local redundancy may be hidden within a system of high total redundancy. Then we may discuss
   - determinability of parameters
   - controllability of errors and the effect of nondetectable errors
   - separability of error sources.

   We will formalize this classification in more detail and discuss the first two items explicitely.
2. How small is too small?
   Most algorithms are controlled by *thresholds* or tolerances to be specified by the developer or the user.
   When referring to observations or parameters, thresholding may be interpreted as hypothesis testing, which allows to derive the thresholds by specifying a significance level and using error propagation. We will not pursue this topic.
   When evaluating, the *design* of the formalization becomes less obvious, e. g. when having a small basis in relative orientation (2D – 2D), small angles in spatial resection (3D – 2D) or small distances between all points in absolute orientation (3D – 3D). In all cases the configuration is close to critical. But then the question arises: how to evaluate *small deviations from a critical configuration or surface*? We will show that a generic and formal answer to this question can be given which is based on the *local* geometry of the design.
3. How to react on deficiencies in the data?
   Regarding the many different models used for calibration and orientation a *generic strategy* should be available.
   Deficiencies in design have to be prevented by proper planning of the mensuration setup influencing the number and position of cameras, the number and the distribution of given control points, the introduction of spatial constraints, etc. Automized techniques for such planning are not far advanced and still require interactive intervention.
   The reaction on *deficiencies in the observations* or the model may rely on the techniques from robust estimation and much more from formalizable experience.
   They depend on various properties of the data and the model:
   - the availability of approximate values $\boldsymbol{\beta}^0$ for the unknown parameters $\boldsymbol{\beta}$.
   - the availability of a direct solution $\boldsymbol{\beta} = \boldsymbol{g}^{-1}(\boldsymbol{l}_s)$ for an $u$-sized subset of the observations.
   - the number and the size of the expected errors.

- the number of the observations and parameters
- the desired efficiency in terms of computing time
- etc.

The next section will collect the necessary tools needed for setting up generic procedures for robust estimation applicable to camera orientation.

# 3 Tools

## 3.1 Quality Insurance

Treating calibration and orientation as an estimation problem allows us to fully exploit the rich arsenal of tools from estimation theory. Regarding the specific problem of data and model errors we specifically need to use the techniques available from robust statistics and regression diagnostics following two different aims (Huber 1991):

- The purpose of *robustness* is to have safeguards against deviations from the assumptions.
- The purpose of *diagnostics* is to find and identify deviations from the assumptions.

**Robustness** There are two levels of robustness, depending on whether the size of errors is small or large. Data or model deviations are small in the case of sufficient linear approximations. This leads to a rule of thumb that small deviations of the approximate values from the true values are deviations less than about 30 % of the values, including all functions of the observations. E. g., it corresponds to the requirement that angular errors to be less than approx. $20^o$.

1. Robustness with respect to *small deviations*.
   The socalled *influence curve* (Hampel et al. 1986), which measures the effect of errors onto the result, may be used to measure the quality of robust procedures in this case. Maximum-likelihood (ML) type, or *M-estimators* are the tools to deal with small deviations.
2. Robustness with respect to *large deviations*.
   The *break down point* (Rousseeuw/Leroy 1987) measuring the maximum allowable percentage in the number of errors while still guaranteeing the estimator to yield results with limited bias, may be used to evaluate the quality of procedures in this case. Estimates with a high break down point, up to 50 %, such as least median squares, are the corresponding tool to handle a large percentage of errors.

Observe, that the effect of *random* errors on the result is not covered by the term robustness. These effects usually are measured by the *precision* of the estimates. The reason for this distinction is that random errors are part of the original model, thus do not represent deviations from the model, and are taken into account by all basic estimators such as least squares or ML-estimators.

We will discuss the use of different robust estimators in section 3.4 and 4, where we especially compare and use their characteristics for achieving a generic strategy.

**Diagnostics** As already indicated above, there are three levels of diagnostics which all refer to small model errors:

1. *Determinability* of parameters or singularities in the estimation process measure the instability of the design with respect to random perturbations.
   Standard deviations or in general covariance matrices are the diagnostic tool to detect such a situation. Due to the small size of the random errors, a linear substitute model derived by linearization, may be used to evaluate such instabilities.
   We will discuss this in detail in section 3.2.
2. *Controllability* of observations and detectability of model errors specify the applicability of hypothesis tests.
   The diagnostic tools are minimum bounds of the size of observational or model errors which can be detected by a test with a certain given probability.
   The *sensitivity* of the result is measured by the effect of nondetectable errors on the result.
   Both tools may be used for planning as they do not depend on the actual measurements.
   The actual influence of the observations of model parameters measured in a leave-one-out fashion may be decisive for the acceptance of an estimate.
   We will discuss these tools in detail in section 3.3.
3. The *locatability* of observational errors or the separability of model errors specify the ability to correctly classify or identify the error causes.
   This can be described in terms of a confusion matrix, like in statistical pattern recognition, the difference being that here the entries of the confusion matrix depend on the expected size of the errors and on the design or configuration.
   The diagnostic tools therefore are lower bounds for observational errors or model errors which are identifiable or separable with a certain probability.
   In section 3.3 we will formally relate separability to controllability especially with respect to *sets* of observational model errors, but will not discuss the notion in detail.

### 3.2 Instabilities of Estimates or "How Small is too Small?"

Instabilities of parameters occur in case the configuration produces some critical manifold (surface) to which the solution belongs. One usually distinguishes (cf. Wrobel 1995):

1. Singularities or critical surfaces of the first kind. Here a complete manifold of the parameters is consistent with the observations.
2. Singularities or critical surfaces of the second kind. Here small deviations in the observations result in large deviations in the parameters.

An example for a singularity of the second kind is the critical cylinder in spatial resection. It may be formulated as a rule: `IF the projection center` $O \in$ `cylinder`$(P_1, P_2, P_3)$ `THEN` $O$ `is not determinable.` Here `cylinder` $(P_1, P_2, P_3)$ indicates the cylinder through the points with axis perpendicular to the plane through the points.

This rule is the result of an analysis using algebraic geometry which, in its generality, is valid in the context of spatial resection and is crisp.

Such algebraic results, however, have some disadvanteges:

- The statements do not contain any information on how to evaluate *deviations from the critical configuration.*
- The statements do not give any hint to *generalize* to other situations. Other problems, e. g. relative orientation, require a separate analysis.
- The statements do not give any *means to evaluate* the orientation even of one image *within a set* of several images to be oriented simultaneously. It may very well be, that in a multi-image setup with a large redundancy the orientation of one of the images cannot be determined due to the presence of the above situation.

Such *hidden instabilities* reveal the limitation of purely algebraic approaches which can only be applied to very restricted situations and cannot be generalized.

Thus techniques based on algebraic geometry cannot be easily transferred into automatic procedures evaluating the stability of an estimate. The solution to this dilemma is based on the observation, that the instabilities are local properties in parameter space and can be fully analysed using the covariance matrix of the parameters. This leads to a shift of the problem. Instead of a deterministic analysis we now are confronted with the problem of evaluating the quality of a covariance matrix. The shift of the problem and its solution goes back to Baarda 1973.

The evaluation method consists of two steps:

1. Specification
   Specifying the user requirements in terms of a socalled *criterion matrix*, say $\boldsymbol{H}$, which gives an upper bound on the desired covariance matrix, corresponding to the desired lowest precision.
2. Comparison
   Checking whether the achieved covariance matrix, say $\boldsymbol{G} = (\boldsymbol{X}^T \boldsymbol{\Sigma}_{ll} \boldsymbol{X})^{-1}$ is better than $\boldsymbol{H}$.

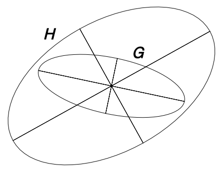
We will discuss this comparison first.

**Comparing Covariance Matrices** The *comparison* of covariance matrices is interpreted as the requirement the standard deviation of an arbitrary function $\boldsymbol{f}$ to be better when calculated with covariance matrix $\boldsymbol{G}$ than with $\boldsymbol{H}$

$$\boldsymbol{G} \leq \boldsymbol{H} \qquad \doteq \qquad \sigma_f^G \leq \sigma_f^H, \text{ with } f = \boldsymbol{e}^T \widehat{\boldsymbol{\beta}}, \text{ for all } \boldsymbol{e} \tag{5}$$

Using error propagation, e. g. $\sigma_f^G = \sqrt{\boldsymbol{e}^T \boldsymbol{G} \boldsymbol{e}}$ this leads to (cf. Fig. 1)

**Fig. 1.** shows the relation $G < H$ between two $2 \times 2$ covariance matrices $G$ and $H$, represented by isolines of constant probability density of the corresponding normal distribution.




$$\boldsymbol{e}^T \boldsymbol{G} \boldsymbol{e} \leq \boldsymbol{e}^T \boldsymbol{H} \boldsymbol{e}, \text{ for all } \boldsymbol{e} \tag{6}$$

or

$$\lambda = \frac{\boldsymbol{e}^T \boldsymbol{G} \boldsymbol{e}}{\boldsymbol{e}^T \boldsymbol{H} \boldsymbol{e}} \leq 1 \tag{7}$$

which requires the determination of the maximum eigenvalue of

$$\boldsymbol{G} \boldsymbol{e} = \lambda \boldsymbol{H} \boldsymbol{e}. \tag{8}$$

The square root $\sqrt{\lambda_{max}}$ indicates the maximum ratio of the actual and the required standard deviation.

This evaluation may be simplified using

$$\boldsymbol{K} = \boldsymbol{H}^{-1/2} \boldsymbol{G} \boldsymbol{H}^{-1/2} \tag{9}$$

$$\lambda = \frac{\boldsymbol{e}^T \boldsymbol{K} \boldsymbol{e}}{\boldsymbol{e}^T \boldsymbol{e}} \leq 1 \tag{10}$$

which is equivalent to

$$\lambda_{\max}(\boldsymbol{K}) \leq 1. \tag{11}$$

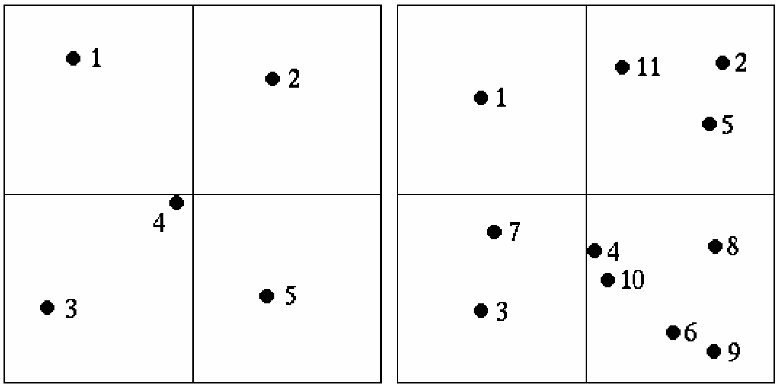
Equation (10) is favorable in case $\boldsymbol{H}$ easily can be diagonalized (cf. the example below).

In order to avoid the rigorous determination of the maximum eigenvalue of $\boldsymbol{K}$, (10) may be replaced by a less tight norm, e. g. by the trace:

$$\lambda_{\max}(\boldsymbol{K}) \leq tr\boldsymbol{K} \leq 1. \tag{12}$$

**Specification of a Criterion Matrix** The *specification* of a criterion matrix can be based on the covariance matrix $\boldsymbol{\Sigma}_{\widehat{\beta}\widehat{\beta}}$ derived from an ideal configuration. This has the advantage that the user can easily interprete the result. In case an ideal configuration cannot be given the criterion matrix $\boldsymbol{H} = \boldsymbol{SRS}$ may be set up by specifying the standard deviations $\sigma_i$, collected in a matrix $\boldsymbol{S} = Diag(\sigma_i)$ and correlations $\rho_{ij}$, collected in a matrix $\boldsymbol{R} = \rho_{ij}$, derived from some theoretical considerations, e. g. interpreting the sequence of projection centres in a navigation problem as stochastic process, where the correlations $\rho_{ij}$ depend only on the time or space difference between points $P_i$ and $P_j$.

**Fig. 2.** shows two sets of image points used for image orientation by spatial resection. Sets of three points may lead to results of different stability as shown in table 1 for three sets of the left configuration (a) (from SCHICKLER [1992]).




**Example** Five image points situated as in Fig. 2 are to be used to estimate the 6 orientation parameters of the image based on given 3D-coordinates with spatial resection (2D – 3D). Due to gross errors in the data, a RANSAC procedure (cf. BOLLES/FISCHLER 1981) is applied randomly selecting 3 points and directly solving for the orientation parameters. The quality of this selection has to be evaluated automatically in order to immediately exclude unstable configurations. The above mentioned technique for evaluating the stability of a configuration is applied.

The criterion matrix is derived from a very stable least squares fit with 4 points symmetrically sitting in the four corners of the image (cf. Appendix). The covariance matrix $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}_{\widehat{\beta}\widehat{\beta}}$ of this configuration, the criterion matrix, is chosen to be

$$\boldsymbol{H} = 16 \cdot \boldsymbol{\Sigma} \tag{13}$$

thus requiring the standard deviations of the orientation parameters within the RANSAC-procedure to be better than 4 times the standard deviation of the ideal configuration. $\boldsymbol{\Sigma}$ is sparse allowing easily, i. e. algebraically, to determine the matrix $\boldsymbol{H}^{-\frac{1}{2}}$ in (10) (cf. Appendix).

For several triplets of points the ratio $\sqrt{\lambda_{max}}$ is given.

| | configuration | $\sqrt{\lambda_{max}}$ |
|---|---|---|
| 1 | 1/2/3 | 0.8 |
| 2 | 2/3/4 | 88.0 |
| 3 | 1/3/4 | 13.2 |

**Table 1.** shows the stability with sets of three points used of spatial resection (cf. Fig. 2a).

The good triangle (1,2,3) obviously leads to sufficiently precise orientation parameters. The second triplet (2,3,4) consists of three nearly collinear points, which obviously is an undesirable configuration. The third triplet (1,3,4) and the projection centre are lying *approximately* on a critical cylinder causing the diagnostic value $\sqrt{\lambda_{max}}$ to be significantly larger than 1., expressing the fact that some function of the orientation parameters in that configuration has a standard deviation being appr. 13 times larger than required. The small triplet (2,5,11) in Fig. 2b also leads to a weak determination of the orientation parameters with a value $\sqrt{\lambda_{max}} \approx 4$.

The method obviously is able to capture various deficiencies in the design of the configuration of an orientation procedure without having to discriminate between different types of instabilities. Such situations also may arise in more complex problems where an algebraic analysis is not possible whereas this method is able to find the instabilities.

When using this method for designing a configuration the eigenvector belonging to the largest eigenvalue gives insight into the most imprecise function of the parameters, which may be used to look for specific stabilization means.

### 3.3 Model Errors or "How Sensitive is the Result?"

The stability of an estimation, specifically an orientation, evaluated by the covariance matrix only takes random perturbations into account. The result, however, may be wrong due to gross errors, e. g. caused by the matching procedure. As well, an oversimplified model may lead to precise but incorrect results. Both error sources, blunders and systematic errors, can only be detected in the case of redundant observations. This is a necessary but - as we will see - not a sufficient condition. Redundancy allows us to perform tests on the validity of the assumed model *without* reference to additional data used during the estimation. Such tests may lead to the detection or even identification of the error source. Of course, the outcome of these tests may be false. Redundancy, however, increases the stability of the solution and the correctness of the outcome of statistical tests. The theory for performing such a test is described in the literature (cf. BAARDA 1967/1968, FÖRSTNER 1987). The structure of that theory, its use in estimation problems and examples from orientation procedures will be given.

**Detectability and Separability of Errors** We first want to discuss the type of evaluation which can be performed depending on the redundancy $r$ of a system.

1. $r = 0$ In the case of no redundancy, one can only evaluate the sensitivity of the result with respect to random errors as shown in the last section. No check of the observations is possible whatsoever. They may remain incorrect without any indication.
2. $r = 1$ In the case of redundancy $r = 1$, a check on the validity of the model is possible. The existence of blunders may be indicated. However, they are not locatable, as a "leave-one-out test" always leads to a valid solution.
3. $r = 2$ A redundancy of $r = 2$ is necessary in order to be able to locate a single blunder. A leave-one-out test generally will be able to find the unique consistent set of observations. Double errors are not locatable, however their existence is usually indicated.
4. $r > 2$ For a larger redundancy, $r - 1 < n/2$, errors are locatable, whereas $r$ errors are only detectable.

The maximum number of detectable errors is $n/2$, i. e. 50 % of the data, as more than $n/2$ observations may mimic a good result. Thus, 50 % is the upper limit for the so-called *breakdown point* of an estimator. The breakdown point of an estimator is the minimum percentage of errors which may cause the estimator to give wrong results, i. e. may lead to a bias of any size. The normal mean has the breakdown point 0, the median 50 %, an indication of it's higher robustness. Practical procedures may be better as they may use specific knowledge about the structure of the problem (cf. the straight line detection procedure by ROTH/LEVINE 1990).

In case of a small percentage ($< 1\%$) of not too large ($\leq 30\%$) gross errors, the detection and location may be based on the residuals

$$\boldsymbol{v} = \boldsymbol{g}(\widehat{\boldsymbol{\beta}}) - \boldsymbol{l} \qquad D(\boldsymbol{y}) = \sigma_0^2 \boldsymbol{Q} = \sigma_0^2 \boldsymbol{P}^{-1} . \tag{14}$$

Using the maximum likelihood estimate

$$\widehat{\boldsymbol{\beta}} = \boldsymbol{\beta}^{(0)} + (\boldsymbol{X}^T \boldsymbol{P} \boldsymbol{X})^{-1} \boldsymbol{X}^T \boldsymbol{P} (\boldsymbol{y} - \boldsymbol{g}(\boldsymbol{\beta}^{(0)}) \tag{15}$$

we can express changes $\Delta \boldsymbol{v}$ of the residuals in terms of changes, thus errors $\Delta \boldsymbol{y}$ of the observations

$$\Delta \boldsymbol{v} = -\boldsymbol{R}\Delta \boldsymbol{y} \tag{16}$$

with the projection matrix

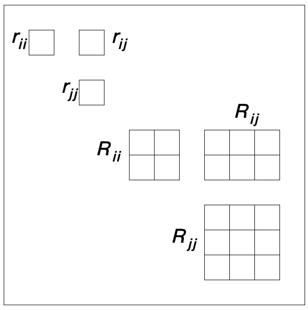
$$\boldsymbol{R} = \boldsymbol{I} - \boldsymbol{U} \tag{17}$$

with the socalled hat-matrix (cf. Huber 1981)

$$\boldsymbol{U} = \boldsymbol{X}(\boldsymbol{X}^T \boldsymbol{P}\boldsymbol{X})^{-1}\boldsymbol{X}^T \boldsymbol{P}\,. \tag{18}$$

(17) is graphically shown in Fig. 3.

**Fig. 3.** shows the four cases for analysing the projection matrix **R** with respect to sensitivity (diagonal matrices) and separability (off-diagonal matrices) for single or groups of observations.




This matrix may be used to analyse the ability of the estimation system to apply selfdiagnosis with respect to errors in the observations, as only effects that can be seen in the residuals are detectable.

We distinguish two levels of evaluation

1. detectability or checkability; and,
2. separability or locatability.

Both evaluation measures may refer to single or groups of observations. Thus we have 4 cases.

1. *Detectability* or checkability rely on the diagonal elements or diagonal submatrices of $\boldsymbol{R}$.
   a) *Single* observational errors can only be detected if the *redundancy numbers*

$$r_i \doteq (\boldsymbol{R})_{ii} > 0\,. \tag{19}$$

The diagonal elements $r_i$ sum up to the total redundancy $r$, i.e. $\sum r_i = r$. This indicates how the redundancy is distributed over the observations. The corresponding test statistics for detecting single errors for given $\sigma_0$ and uncorrelated observations is

$$z_i = \frac{-v_i}{\sigma_0}\sqrt{\frac{p_i}{r_i}} \sim N(0\,,\,1) \tag{20}$$

b) *Groups* of $n_i$ observation can only be detected if the corresponding $n_i \times n_i$ submatrix

$$\| R_{ii} \| > 0 \tag{21}$$

of $\boldsymbol{R}$ is nonsingular. Otherwise a special combination of observational errors may have no influence on the residuals. The corresponding test statistic is

$$T_i = \frac{1}{\sigma_0}\sqrt{\frac{\boldsymbol{v}_i^T \boldsymbol{R}_{ii} \boldsymbol{Q}_{ii} \boldsymbol{v}_i}{n_i}} \sim \sqrt{F_{n_i,\infty}} \tag{22}$$

which reduces to (20). The observations may be correlated within the group, but must be uncorrelated to the others. $\sqrt{\boldsymbol{F}_{n_i,m}}$ denotes the distribution of the square root of a random variable being $F_{n_i,m}$-distributed.

2. *Separability* or locatability in addition to the diagonal elements of $\boldsymbol{R}$ rely on the off diagonals.
   a) The separability of two *single gross errors* evaluates the likelihood to correctly locate an error, i. e. to make a correct decision when testing both. The decisive measure is the correlation coefficient of the test statistics (20) which is

$$\rho_{ij} = \frac{r_{ij}}{\sqrt{r_{ii} \cdot r_{jj}}}. \tag{23}$$

   Tables for erroneous decisions when locating errors are given by FÖRSTNER 1983.
   Correlation coefficients below 0.9 can be accepted since the probability of making a false decision even for small errors remains below 15 % [1].
   b) The separability of two *groups of observations* $\boldsymbol{l}_i$ and $\boldsymbol{l}_j$ depends on the maximum value

$$\rho_{ij}^2 = \lambda_{max} \boldsymbol{M}_{ij} \tag{24}$$

   of the $n_i \times n_j$ matrix

[1] Precisely stated: If the larger of the two test statistics $|z_i|$ and $|z_j|$ in (20) is chosen to indicate the erroneous observation with its critical value 3.29, corresponding to a significance level of 99.9 %, and a single error can be detected with a probability higher than 80 %, then the probability of making a wrong decision between $l_i$ and $l_j$ is approximately 13 %.

$$\boldsymbol{M}_{ij} = \boldsymbol{R}_{ij}\boldsymbol{R}_{jj}^{-1}\boldsymbol{R}_{ji}\boldsymbol{R}_{ii}^{-1} \tag{25}$$

which for single observations reduces to (23).
No statistical interpretation is available due to the complexity of the corresponding distribution.

**Example:** *Detectability of Errors*

Relative orientation with 6 corresponding points yields a redundancy of $r = 6 - 5 = 1$. If the images are parallel to the basis and the points are situated symmetrically as shown in Fig. 4 then the diagonal elements $r_i$ are 1/12 for points $i = 1, 2, 5$ and 6 and 1/3 for points 3 and 4.

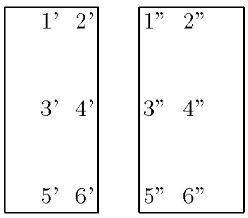



**Fig. 4.** Numbering of 6 points in a stereo pair.

Obviously errors are hardly detectable if they occur in point pairs 1, 2, 5 or 6. In all cases no location of the false matches is possible as $r = 1$. □

**Example:** *Separability of Errors*

Spatial resection with 4 points symmetrically to the principle point is known to yield highly correlated orientation parameters. Depending on the viewing angle $\alpha$, the correlation between the rotation $\omega$ ($x$-axis) and the coordinate $y_0$ of the projection centre, and between the rotation $\varphi$ ($y$-axis) and the coordinate $x_0$ is (cf. Appendix)

$$| \rho | = \frac{1}{\sqrt{1 + \sin^4 \frac{\alpha}{2}}} \tag{26}$$

For a CCD-camera with a focal length of $f = 50$ mm and sensor size of $5 \times 5mm^2$ , $\alpha/2 = 1/20$ thus $| \rho | = 0.999997$. For an aerial camera RMK 15/23 with a focal length of 15 cm and image size of 23 cm, $\alpha/2 = 2/3$, thus $| \rho | = 0.914$.

Thus testing the orientation parameters $\omega$, $\varphi$, $x_0$ and $y_0$ may easily lead to incorrect decisions for CCD-cameras when testing their significance, whereas errors in these parameters are detectable. □

**Sensitivity of the Estimates** In spite of testing for blunders, errors may remain undetected and influence the resulting estimate. The *sensitivity* of the result is often the only information one needs for evaluation. One may determine an upper limit for the influence of a group of observations onto the result.

The influence $\Delta_i f(\widehat{\boldsymbol{\beta}})$ on a function $f(\widehat{\boldsymbol{\beta}})$ of the unknown parameters caused by leaving out a group $\boldsymbol{y}_i$ of observation is limited:

$$\Delta_i f(\widehat{\boldsymbol{\beta}}) \leq \Delta_i f_{max}(\widehat{\boldsymbol{\beta}}) \tag{27}$$

with (cf. Förstner 1992)

$$\Delta_i f_{max}(\widehat{\boldsymbol{\beta}}) = T_i \cdot \mu_i \cdot \sigma_{f(\beta)} \cdot \sqrt{n_i} \tag{28}$$

where $n_i$ is the size of the group, $\sigma_{f(\beta)}$ the standard deviation of the function $f(\beta)$ is derivable by error propagation measuring the precision of the result, $T_i$ of the test statistics (22), measuring the quality of the observation group and the geometry factor

$$\mu_i = \lambda_{max}\{(\boldsymbol{\Sigma}_{xx}^{(i)} - \boldsymbol{\Sigma}_{xx})\boldsymbol{\Sigma}_{xx}^{-1}\} \tag{29}$$

evaluating the mensuration design. The value $\mu_i$ explicitely measures the loss in precision, i. e. the normalized increase $\boldsymbol{\Sigma}_{\beta\beta}^{(i)} - \boldsymbol{\Sigma}_{\beta\beta}$ of variance of the result when leaving out the $i$−th group $\boldsymbol{l}_i$ of observations.

For a single observation it reduces to

$$\mu_i = \frac{1 - r_i}{r_i} \tag{30}$$

with the diagonal elements $r_{ii}$ of $\boldsymbol{R}$ (cf. (17)).

The value $\Delta_i f_{max}(\boldsymbol{\beta})$ (28) measures the *empirical sensitivity* of the estimate with respect to blunders e. g. matching errors in groups $\boldsymbol{l}_i$; empirical, as it depends on the actual observations via $T_i$.

If $T_i$ is replaced by a constant $\delta_0$, indicating the minimum detectable (normalized) error, we obtain the *theoretical sensitivity*

$$\Delta_{0i} f(\widehat{\boldsymbol{\beta}}) \leq \Delta_{oi} f_{max}(\widehat{\boldsymbol{\beta}}) \tag{31}$$

with

$$\Delta_{0i} f_{max}(\widehat{\boldsymbol{\beta}}) = \delta_0 \cdot \mu_i \cdot \sigma_{f(\beta)} \cdot \sqrt{n_i}\,. \tag{32}$$

It may be used for planning purposes since it does not depend on actual observations and can therefore be determined in advance. $\delta_0$ is usually chosen to be larger than the critical value $k$ for $T_i$, e.g. $\delta_0 = 1.5k$ or $\delta_0 = 2k$ and can be linked to the required power of the test (cf. Baarda 1967/1968, Förstner 1987).

Observe that both sensitivity values contain the product of terms representing different causes. This e. g. allows to sacrifice precision, thus increasing standard deviation $\sigma_f$ by paying more for leaving a larger redundancy and lowering the geometric factor $\mu_i$ for all observations or vice versa.

**Example:** *Sensitivity Analysis*

This example shows the power of this type of sensitivity analysis for evaluating the success of an automatic procedure for determining the exterior orientation of an image, i. e. the extrinsic parameters of the camera (SCHICKLER 1992, cf. SESTER/FÖRSTNER 1989). It is based on matching 2D-line segments in the image with 3D-line segments of a set of known objects, mainly being buildings represented by a set of line segments. The aerial images used here usually contain 5-10 such objects which are more or less well-distributed over the field of view.

The sensitivity analysis may be used to evaluate the quality of the orientation with respect to

a) matching errors of *individual line segments*; and,
b) matching errors of complete *sets* of line segments, representing one *object*.

The reason for this distinction is that both errors may occur; the first one being very common, the second one (whole sets of line segments) within the clustering procedure performed for each object individually.

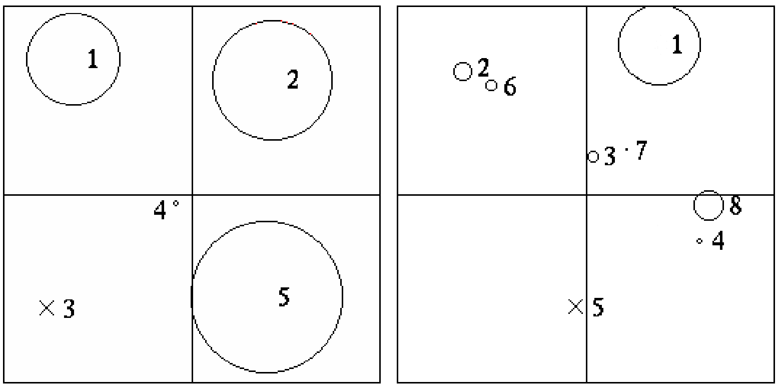
a) Matching of *individual* 2D image line segments to 3D object line segments. We have to deal with groups of 4 observations, namely the 4 coordinates representing the start and end point of the line segments. The $4 \times 4$ covariance matrix $\boldsymbol{\Sigma}_{l_i l_i}$ of this group also contains the correlations between the coordinates, which may be derived during the edge extraction process. We use a similar approach as DERICHE/FAUGERAS 1990 and FÖRSTNER 1992 for representing the uncertainty of the line segments.
   A typical result, as given in the following table, can be summarized in two statements:
   1. Empirical sensitivity: The maximum occurs at edge #10. The result may change up to 0.82 it's standard deviation if line segment #10 would be left out, which is fully acceptable.
   2. Theoretical sensitivity: The maximum occurs at edge #21. The result may change up to 4.42 times its standard deviation if a matching error remains undetected, which is at the limit of being acceptable.

   Thus, the result appears to be acceptable with respect to the redundancy in the estimations.
b) Match of a *set* of 2D image line segments to 3D object line segments.
   Let us assume the $m$ sets of segments to be matched, have $k_i, i = 1, \cdots, m$ line segments each, and we have to fear a matching error for a complete set. Then the sensitivity analysis has to be based on sets of $4 \times m_i$ coordinates for the $m_i$ line segments.
   Figures 5a and 5b show the position of the sets within two aerial images ($c = 15$ cm) to be oriented.
   In Fig. 5a, one of the five sets, namely #3 was not matchable, leaving the spatial resection with 4 objects in the 3 other corners and in the middle of the image. The circles around these "four points" have a radius proportional

**Table 2.** shows the empirical and the theoretical sensitivity of the result of an orientation with straight edge segments.

| Edge # | Empirical $\Delta_i f/\sigma_f$ | Theoretical $\Delta_{0i} f/\sigma_f$ |
|---|---|---|
| 4 | 0.07 | 2.62 |
| 5 | 0.65 | 1.51 |
| 8 | 0.50 | 3.44 |
| 9 | 0.80 | 3.13 |
| 10 | 0.82 | 2.81 |
| ⋮ | ⋮ | ⋮ |
| 21 | 0.68 | 4.42 |
| ⋮ | ⋮ | ⋮ |

**Fig. 5.** shows two sets of image points used for image orientation by spatial resection. The radii of the circles indicate the theoretical sensitivity, i. e. the amount the result might change if the point (set of straight line segments) would be wrong without notice. In Fig. a (left) the point #3 has been detected to be wrong, thus only 4 points are left for spatial resection, in Fig. b (right) point # 5 has been detected to be wrong.




to $\delta_{0i} = \Delta_{0i} f_{max}/\sigma_f$ and indicate how sensitive the orientation is with respect to nondetectable errors within the clustering procedure. Because the geometry factor $\mu$ (29) is dominant, the circles indicate how the precision deteriorates if one of the 4 sets is left out:

**without set 4:** the three others 1, 2 and 5 form a well-shaped triangle, and thus guarantee good precision.

**without set 2:** the three others 1, 4 and 5 nearly sit on a straight line leading to a highly unstable solution (near to singularity of first type).

**without set 1:** the three others, 2, 4 and 5, form a well-shaped triangle.

However, because the plane going through the sets is nearly parallel to the image plane, the projection centre closely has near to the *critical cylinder*.

**leaving out set 5:** also leads to a nearly singular situation.

The situation with 8 sets in Fig. 5b shows a more irregular distribution. Since set 5 was not matched, set 1 is most influential in the orientation, but less than sets 1, 2 and 5 in the case of Fig. 5a.

□

Observe that this analysis is based on values which have a very precise geometric meaning. This allows for an easy definition of thresholds, even if one is not aquainted with the underlying theory. As well, a clear comparison between different configurations is possible even for different types of tasks. Because the evaluation refers to the final parameters, it also may be used when fusing different type of observations. As model knowledge may be formalized in a Bayesian manner, the effect of prior information onto the result of an orientation may also be analysed.

Summarizing the evaluation of the design using the comparison of the covariance matrix of the parameters with a criterion matrix and using the different measures for the sensitivity has several distinct properties:

- it is a general concept
- it works for all types of critical surfaces and solves the problem of critical areas, thus also in case the configuration of observations is far or close to a critical surface
- it works with all problems of estimation
- it may detect hidden singularities
- it also works in the complex situation where observations of different types are mixed (points, lines, circles, ...) or in the context of sensor fusion where also physical measurements (force, acceleration, ...) are involved
- it is related to a task, thus explicitely depends on user requirements. This enables to argue backwards and optimize the design.
- it provides measures which are easily interpretable.

### 3.4 Robust Estimation or "How to React on Blunders"

The last section clearly demonstrated that enough tools are available to evaluate the result of estimation procedures with respect to a variety of deficiencies. These tools are sufficient for proving a result to be acceptable. They, however, give no hint as to how to reach an acceptable result with respect to errors in the data and weaknesses in the design.

This section wants to collect the techniques from robust statistics useful for the efficient elimination or compensation of outliers in the data with the aim of adapting the data to the presumed model. The planning of the mensuration design is much more difficult and lacks enough theoretical basis and is therefore not discussed here.

Eliminating blunders is a difficult problem:

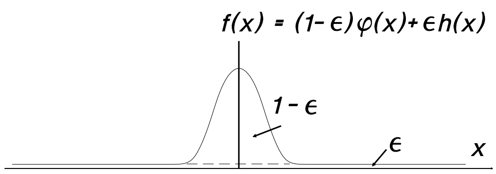
- It is NP-complete: given $n$ observations there are up to $2^n$ hypotheses for sets of good and bad values (the power set of $n$ observations), making an exhaustive search for the optimized solution obsolete except for problems with few observations.
- The non-linearity of most estimation problems, particularly orientation problems, prevents generic simplification for obtaining suboptimal solutions.
- All variations of "Murphy's Law" occur:
  - outliers cluster and support each other,
  - outliers mimic good results,
  - outliers hide behind configuration defects,
  - outliers do not show their causes, making proper modelling difficult or impossible,
  - outliers make themselves indistinguishable from other deficiencies in the model, like systematic errors.

Thus many methods for robust estimation have been developed. Most of them assume the model of a mixed distribution of the residuals $v_i$ ($f$ denoting a density function here):

$$f(v_i/\sigma_i) = (1-\varepsilon)\phi(v_i/\sigma_i) + \varepsilon h(v_i/\sigma_i) \tag{33}$$

with $100\varepsilon\%$ outliers having broad distribution $h(\beta)$ and $100(1-\varepsilon)\%$ good observations following a well-behaved distribution $\phi$, usually a Gaussian. Maximizing $f(\boldsymbol{\beta} \mid \boldsymbol{l})$ or minimizing $-\log f(\boldsymbol{\beta} \mid \boldsymbol{l})$ for the given data $\boldsymbol{l}$, possibly including prior knowledge of the unknowns $\boldsymbol{\beta}$, explicitely or implicitely is used as optimization criterion.

**Fig. 6.** shows the distribution of the errors being a mixture between good and bad ones.



The procedures, however, significantly differ in strategy for finding the optimum or a suboptimum. We selected four procedures which seem to be representative in order to come to an evaluation which will be the basis for the generic strategy discussed in section 4. These techniques for robust estimation are

1. complete search,
2. Random Sample Consensus (RANSAC cf. BOLLES/FISCHLER 1981),
3. clustering, and
4. ML-type-estimation (HUBER 1981, HAMPEL/ET AL. 1986).

Their feasibility and efficiency heavily depend on a number of characteristic features of the estimation problem to be solved.

a) *Invertibility of the Functional Model*
We basically used a set $\boldsymbol{l}_s$ of at least $u$ observation to uniquely specify the unknown parameters $\boldsymbol{\beta}$. The direct determination of $\boldsymbol{\beta}$ from a subset $\boldsymbol{l}_s$ requires $\boldsymbol{g}$ to be invertible: $\boldsymbol{\beta}(\boldsymbol{l}_s) = \boldsymbol{g}^{-1}(\boldsymbol{l}_s)$ thus $\boldsymbol{g}^{-1}$ has to be representable algebraically.

b) *Existence and Quality of Approximate Values*
In case $\boldsymbol{g}(l)$ is not invertible, we need approximate values for $\boldsymbol{\beta}$ in order to solve $\boldsymbol{\beta} = \boldsymbol{g}^{-1}(\boldsymbol{l})$ by *some* iterative scheme. The quality of the approximate values directly influences the number of iterations. The knowledge of good approximate values in all cases may drastically reduce the complexity of the procedures.

c) *Percentage of Gross Errors*
The percentage of gross errors may range from $< 1\%$, specifically in large data sets derived automatically, up to more than 75%, e. g. in matching problems. Not all procedures can cope with any percentage of errors, some are especially suited for problems with high outlier percentages. ML-type-estimation procedures can handle data with a moderate number of errors, up to $10 - 20\%$ say.

d) *Size of Gross Errors*
Only few procedures can work for any size of gross errors. Large gross errors may lead to leverage points, i. e. to locally weak geometry, and such errors may not be detectable at all. If one relates the size of the errors to the size of the observed value, then errors less than one unit usually are detectable by all procedures.

e) *Relative Redundancy*
The relative redundancy measured by the redundancy numbers $r_i$ (cf. eq. (19)) influences the detectability of errors. The theoretical results from robust statistics, especially with reference to ML-type-estimation, are only valid for relative redundancies above 0.8, i. e. when the number of observations is larger than 5 times the number of unknown parameters.

f) *Number of Unknowns*
The number of unknowns directly influences the algorithmic complexity.

The four procedures can now easily be characterized.

1. *Complete Search*

   Complete search checks all, i. e. up to $2^n$, possible configurations of good and bad observations to find the optimum solution. The optimization function obviously should contain a cost-term for bad observations in order not to select a minimum of $\mu$ observations yielding residuals $e_i = 0$, or the best set of $\mu + 1$ observations allowing to estimate $\widehat{\sigma}_0^2$ with only one redundant observation. Such a penalty may be derived using the principle of minimum description length, thus relying on the negative logarithm mixed distribution (cf. Fig. 6).

   Obviously complete search is only feasible for a small number $n$ of observations, a small redundancy $r$ or in case the maximum number $l_{\max}$ of expected errors is small, as the number of possibilities is

   $$\sum_{k=0}^{\min(r-1, l_{\max})} \binom{n}{k} < 2^n \,. \tag{34}$$

   Implementation requires either approximate values or the invertibility of the model using an iterative or a direct solution technique.

2. *Random Sample Consensus (RANSAC)*

   Random Sample Consensus relies on the fact that the likelihood of hitting a good configuration by randomly choosing a set of observations is large. This probability of finding at least one good set of observations in $t$ trials is $1 - (1 - (1 - \varepsilon)^u)^t$ where $u$ is the number of trials and $\varepsilon$ the percentage of errors. E. g. for $u = 3$, (spatial resection, fitting circle in the plane) and an outlier rate of 50 % at least $t = 23$ trials are necessary, if this probability should be larger than 95 %.

   Again, the technique requires approximate values or the invertibility of the model and is only suited for small $u$.

3. *Clustering*

   Clustering consists of determining the probability density function $f_\beta(\boldsymbol{y})$ under the assumption that the data represent the complete sample. The mode, i. e. the maximum, of $f_\beta(\boldsymbol{\beta} \mid \boldsymbol{l})$ is used as an estimate. This is approximated by $f_\beta(\boldsymbol{\beta} \mid \boldsymbol{l}) \approx \sum_i f_\beta(\boldsymbol{\beta} \mid \boldsymbol{l}_s^{(i)})$ where the sum is taken over all or at least a large enough set of subsets $\boldsymbol{l}_s$ of $u$ observations, implicitely assuming these subsets to be independent.

   The Hough-Transformation is a classical example of this technique. Stockman 1987 discusses the technique in the context of pose determination, thus for determining the mutual orientation between an image and a model.

   Clustering is recommendable for problems with few unknowns, high percentage of gross errors and in cases in which enough data can be expected to support the solution (high relative redundancy).

4. *Maximum-likelihood-type Estimation*

   Maximum-likelihood-type estimation is based on an iterative scheme. Usually the method of modified weights is used showing the close relation to the classical ML-estimation, where the observations are assumed to be Gaussian distributed. Instead of minimizing $\sum(e_i/\sigma_i)^2$, the sum of a less increasing function $\rho(e_i/\sigma_i)$ is minimized. This can be shown to be equivalent to iteratively weighting down the observations using the weight function $w(\beta) = \rho'(\beta)/\beta$. For convex and symmetric $\rho$, bounded and monotone decreasing $w(\beta)(\beta > 0)$ and a linear model uniqueness of the solution is guaranteed (Huber 1991). Since the influence function $\rho'(\beta)$ (Hampel/et al. 1986) stays strictly positive in this case, indicating large errors still influencing the result, non-convex functions $\rho$ are used.

Most orientation problems are nonlinear and the influence of large errors should be eliminated, thus approximate values are required when using this ML-type estimation. Further requirements are: moderate sized errors, small percentage of errors and homogeneous design, i. e. large enough local redundancy (no leverage points). The advantage of this technique is its favorable computational complexity being $O(u^3 + nu^2)$ in the worst case allowing to be used also for large $u$ where sparse techniques may be applied to further reduce complexity.

Without discussing the individual techniques for robust estimation in detail, which would uncover a number of variations and modifications necessary for implementation, the techniques obviously are only applicable under certain - more or less precisely known - conditions. Moreover, specific properties both of the techniques and of the problem to be solved suggest the development of heuristic rules for the application of the various techniques leading to a generic strategy for using robust techniques, which will be discussed in the final section.

## 4 Generic Estimation Procedures

Generic estimation procedures need to choose the technique optimal for the problem concerned and be able to evaluate their performance as far as possible. This section discusses a first step in formalizing strategic knowledge and the mutual role of robust estimates and diagnostic tools.

### 4.1 Rules for Choosing Robust Estimation Techniques

The qualitative knowledge about the four robust estimation techniques discussed in the previous section is collected in table 3. It shows the degree of recommendation for each technique dependent on the 8 criteria. These criteria refer to:

- necessary prerequisite (approximate values, direct solution);
- likelihood of success (number of observation, reliability, size and percentage of errors); and,
- computational complexity (number of parameters, speed requirements).

We distinguish 4 degrees of recommendation:

**Table 3.** Shows the properties of four techniques for robust estimation

| | Complete Search | | | RANSAC | | | Clustering | | | ML-type Estimation | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | vg | b | i | vg | b | i | vg | b | i | vg | b | i |
| approximate values | | | – | | | – | + | | – | + | | – |
| direct solution | | – | – | + | – | – | + | – | – | | | |
| many observations | – | + | | + | | | + | – | | + | | |
| few parameters | + | – | | + | – | | + | – | | | | |
| high reliability | | | | + | | | | | | + | – | |
| large errors | + | | | + | | | + | | | – | + | |
| high error rate | + | | | | | | + | | | – | + | |
| speed unimportant | + | – | | | | | + | – | | – | | |

vg = very good (and)
b = bad (or)
i = impossible (and)
(possible = not(impossible))
+ = feature required
– = ¬ feature required

- "very good". In case all indicated criteria are fulfilled ("and"); the technique can exploit its power and usually is best.
- "good". In case none of the criteria for "bad" is fulfilled; the technique works "not bad".
- "bad". In case one of the indicated criteria is fulfilled ("or"); the technique shows unfavorable properties, so is unreliable or too costly.
- "impossible". In case all indicated criteria are fulfilled ("and"); the technique cannot be used.

This knowledge can easily be put into rules, e. g. using PROLOG, together with a few additional rules for qualitative reasoning, e. g. `very-recommendable(X): - good(X), possible(X)` or `impossible(X): - not (impossible(X))`. This allows for the automatic selection of the robust estimation procedure which fits best to the problem at hand.

**Example** The determination of the extrinsic parameter of a camera orientation using sets of straight line segments, already mentioned above (example on sensitivity analysis), is performed in several steps.

Step 1: Estimation of the approximate position (2 parameters) of the projected model of each set in the image in order to obtain a prelimenary set of candidate matches between image and model segments. A sample dialog is given below:

```
Please characterize your problem:
Answer with y=yes, n=no, -=do not know, ?=help

Does a direct solution exist         (Default=n)  ?   y
Are approximate values available     (Default=n)  ?   y
Do you have many observations        (Default=n)  ?   y
Are there many unknown parameters    (Default=n)  ?   n
Is the percentage of errors large    (Default=y)  ?   y
Do you expect large blunders         (Default=y)  ?   y
Is computational speed essential     (Default=y)  ?   n


Strongly recommended:
     clustering
     ransac
Not recommended:
     ml type estimation
     complete search
```

Clustering and RANSAC are strongly recommended. ML-type estimation is not recommended as large errors are to be expected. Complete search is not recommended as the number of observations is large.

Step 2 Estimation of good approximate values for the 6 orientation parameters based on the reference points for each object. Thus only few (point) observations are available. The decision is shown in table 4. As computational speed is made essential in this step and the percentage of errors is large (e. g. 3 out of 8 points) only RANSAC is recommendable.

Step 3 The final cleaning of the observations again refers to the line segments. The system does not have access to a direct solution (e. g. by Horaud et al. 1989) and is required to be fast. Therefore ML-type estimation is highly recommendable.

Obviously the qualitative reasoning may be made more precise:

- The number of observations (few, many) is actually known in a special situation. It influences the density of the cells in clustering, the relative redundancy specifically the homogeneity of the design and the likelihood of finding a good set in RANSAC.
- The number of unknowns (few, many) also is known in a specific situation and can be used to predict the computational effort quite precisely.
- The homogeneity of the design usually can be approximated in case the number of observations is much higher than the number of the unknowns, which actually was implemented as a rule in the above mentioned PROLOG program.
- The size and the percentage of the errors to be expected can be predicted from previous data sets and information which vision algorithms should report for learning their performance.

**Table 4.** Shows the decision of a PROLOG program for the selection of the appropriate robust estimation technique.

| | Step 1[1)] | Step 2 | Step 3 |
|---|---|---|---|
| direct solution | yes | yes | no |
| approximate values | yes | yes | yes |
| many observations | yes | no | yes |
| many unknowns | no | no | no |
| many errors | yes | yes | no |
| large errors | yes | no | no |
| speed essential | no | yes | yes |
| very recommendable | clustering RANSAC | - | ML-type |
| recommendable | clustering RANSAC | RANSAC | ML-type RANSAC |
| not recommendable | ML-type complete search | all except RANSAC | clustering complete search |

[1)] cf. sample dialog

– The required speed usually can be derived from the specification of the application and be quite rigorously related to the available resources. An example for such a performance prediction in the context of recognition tasks is given by Chen/Mulgaonkar 1990.

The final goal of a formalization of tool selection would be to leave the choice of the appropriate estimation procedure to the program, which of course requires the standardization of the input/output relation for the procedures.

### 4.2 Integrating Robust Estimation and Diagnosis

For achieving results which are robust with respect to model errors an integration is necessary which exploits the diagnostic tools and the robust procedures. Diagnostic tools do not influence the estimates and robust estimates only work if the design is homogeneous, requiring a rigorous diagnosis.

Therefore, three steps are necessary:

1. The mensuration design has to be planned in order to guarantee that model errors are detectable and undetectable model errors have only acceptable influence on the result. Diagnostic tools are available for all type of model errors; gross errors, systematic errors or errors in distribution. Strategies for planning, however, are poorly formalized and up to now require at least some interactive effort.
2. Robust estimation techniques may then be used to find an optimal or at least a good estimate for the unknown parameters. This actually can be viewed as

an hypothesis generation about the quality (good, bad) of the observations, which is obvious in matching problems which use robust techniques. Any technique may be used which leads to good hypothesis.

3. In the final step the parameters are optimally estimated, e. g. using ML-type estimation based on the decision made in the previous step. The result can then be rigorously checked if it meets the requirements set up in the planning phase. Thus here again the diagnostic tools are used. The result of this analysis provides an objective self-diagnosis which may then be reported to the system in which the estimation procedure is embedded.

The overall quality of the estimation procedure is its ability to correctly predict its own performance, which of course can only be checked empirically (Schickler 1992).

**Example** Table 5 summarizes the result of 48 image orientations.

The total number of correct and false decisions of the selfdiagnosis is split into the cases where the images contained 6 or more points, i. e. sets of straight line segments, and cases with 5 or less points. An orientation was reported as correct if the empirical and the theoretical sensitivity factors $\Delta_i f/\sigma_f$ and $\Delta_{0i} f/\sigma_f$ (cf. eq. (28) and (31)) and the standard deviations of the result were acceptable ($\lambda_{max} < 1$ cf. eq. (12)).

**Table 5.** shows the result of an extensive test of orienting 48 aerial images, # of cases, in brackets: for $\geq 6$ points/image and for $\leq 5$ points/image.

| | | report of selfdiagnosis | |
|---|---|---|---|
| | | correct | wrong |
| reality | correct | 46(39/7) (correct decision) | 0(0/0) (false positives) |
| | wrong | 1(0/1) (false positives) | 1(0/1) (correct decisions) |

46 out of 48 orientations were correct and this was reported by the self-diagnosis. In one case the orientation was incorrect, which was detected by the analysis. This appeared in an orientation with only 4 points, thus only one redundant point. Therefore altogether in 47 out of 48, i. e. in 98 % of all orientations the system made a correct decision. In about half of the cases (22 out of 46) the RANSAC procedure was able to identify errors which occured during the clustering and correct the result of the clustering, which was repeated with this a priori knowledge.

One orientation failed without being noticed by the system, which corresponds to 2 % false positives. This was an orientation with only 5 points.

The orientation of the 48 images was based on 362 clusterings of model and image line segments. 309, thus 85 %, were correct. As the errors in clustering are either completely wrong and therefore eliminated from the further processing or are wrong by a small amount, it is quite likely that 2 clusterings are incorrect by only a small amount, which may not be detectable by the RANSAC or the robust ML-type estimation, mimicing a good orientation. Therefore the existence of one false positive is fully acceptable.

The result achieved in this test is a clear reason to require at least 6 points, i. e. sets of straight edges, for a reliable orientation in this application. As can be seen in the table, all 39 orientations not only could correctly be handled by the automatic system, but actually lead to correct orientation parameters.

This example reveals the diagnostic tools to be extremely valuable for a final evaluation of an automatic procedure containing robust estimation procedures as parts.

## 5 Conclusions

The goal of this paper was to collect the tools from robust statistics and diagnostics necessary for building fully automatic image analysis procedures, specifically orientation procedures. The theory available seems to be sufficient for achieving a high degree of selfdiagnosis and for implementing generic strategies based on knowledge about the specific properties of the different estimation techniques.

In all cases the general strategy for achieving results of high quality consists of four steps:

1. Planning the mensuration configuration using the diagnostic tools for precision and sensitivity garanteeing robust estimation techniques in step 3 to be applicable and the result to be evaluated internally.
2. Mensuration according to the planned configuration which itself gives a clear indication of its quality.
3. Robust estimation with any of the available techniques leading to a hypothesis of a good result.
4. Final evaluation based on the result of an optimal estimation of the parameters and check in how far the quality intended in the planning stage actually is reached.

The examples give clear indication that these tools can be used to advantage even in comparatively complex situations.

There are still some questions open:

- The analysis of the precision of the result is based on the comparison of the actual with the required covariance matrix. Generating meaningful criterion matrices requires proper *modelling of the user needs* making the specification of the accuracy requirements a nontrivial problem.

- The effect of systematic errors (biases in the model) can be analysed with the same techniques. However, the *search space for identifying undetectable systematic errors is large*, due to the unknown interference between the different causes for such errors.
- The planning of experiments may be based on the techniques collected in this paper. Automating the planning, as it may occur in active vision, however, requires the development of strong *strategies* for finding optimal or at least satisfying observation *configurations*.
- The selfdiagnosis, based on the precision and sensitivity analysis, does not give indications on the *probability* of the result to be correct. This would enable the calling routine to react in a more specific manner or to use this probability for further inference.

It would, however, be of great value if all orientation procedures would offer at least the available measures for making a proper selfdiagnosis in order to objectify the quality of the very first steps within image analysis.

## A Algebraic expression for the normal equations of spatial resection with four parts in symmetric position.

Let points $P_i$ with coordinates $(x_i, y_i, z_i), i = 1, \ldots, n$ in the camera system be given and observed in the image. The linearized observation equations for the image coordinates $(x', y')$ depending on the 6 orientation parameters, namely the rotation angles $\omega, \varphi, \kappa$ and the position $(x_0, y_0, z_0)$ of the projection centre can be expressed as

$$dx'_i = -\frac{c}{H} z_0 dx_0 - \frac{x'_i}{H} z_0 dx_0 - \frac{x'_i y'_i}{c} d\omega + c(1 + \frac{x'^2_i}{c^2}) d\phi + y'_i d\kappa \tag{35}$$

$$dy'_i = -\frac{c}{H} z_0 dy_0 - \frac{y'_i}{H} z_0 dy_0 - c(1 + \frac{y'^2_i}{c^2}) d\omega + \frac{x'_i y'_i}{c} d\varphi - x'_i d\kappa \tag{36}$$

valid for each image point $P'(x'_i, y'_i)$. $c$ is the camera constant.

In case $n = 4$ image points lie in symmetric position $(\pm d, \pm d)$ in the image (cf. Fig. 7) and the $z$-coordinates of the points $P_i(x_i, y_i, z_i)$ in the coordinate system of the camera are equal to $H = z_i$, we can collect the coefficients of the $8 \times 6$ matrix $\boldsymbol{X}$ as in the table below.

The algebraic expression for the normal equation matrix $\boldsymbol{N} = \boldsymbol{X}^T \boldsymbol{P} \boldsymbol{X}$ assuming weights 1 for the observations is given by

$$\boldsymbol{N} = \begin{pmatrix} 4\frac{c^2}{H^2} & 0 & 0 & 0 & -4\frac{e^2}{H} & 0 \\ 0 & 4\frac{c^2}{H^2} & 0 & 4\frac{e^2}{H} & 0 & 0 \\ 0 & 0 & 6\frac{d^2}{H^2} & 0 & 0 & 0 \\ 0 & 4\frac{e^2}{H} & 0 & 4\frac{e^4}{c^2}(1 + \frac{d^4}{e^4}) & 0 & 0 \\ -4\frac{e^2}{H} & 0 & 0 & 0 & 4\frac{e^4}{c^2}(1 + \frac{d^4}{e^4}) & 0 \\ 0 & 0 & 0 & 0 & 0 & 8d^2 \end{pmatrix} \tag{37}$$

**Fig. 7.** shows the normalized situation for spatial resection used as a reference for precision.

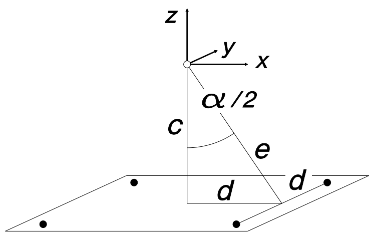



**Table 6.**

| $i$ | | $dx_0$ | $dy_0$ | $dz_0$ | $d\omega$ | $d\phi$ | $d\kappa$ |
|---|---|---|---|---|---|---|---|
| 1 | $x' = +d$ | $-\frac{c}{H}$ | 0 | $-\frac{d}{H}$ | $-\frac{d^2}{c}$ | $\frac{e^2}{c}$ | $+d$ |
| | $y' = +d$ | 0 | $-\frac{c}{H}$ | $-\frac{d}{H}$ | $-\frac{e^2}{c}$ | $\frac{d^2}{c}$ | $-d$ |
| 2 | $x' = -d$ | $-\frac{c}{H}$ | 0 | $+\frac{d}{H}$ | $+\frac{d^2}{c}$ | $\frac{e^2}{c}$ | $-d$ |
| | $y' = +d$ | 0 | $-\frac{c}{H}$ | $-\frac{d}{H}$ | $-\frac{e^2}{c}$ | $\frac{d^2}{c}$ | $-d$ |
| 3 | $x' = -d$ | $-\frac{c}{H}$ | 0 | $+\frac{d}{H}$ | $-\frac{d^2}{c}$ | $\frac{e^2}{c}$ | $-d$ |
| | $y' = -d$ | 0 | $-\frac{c}{H}$ | $+\frac{d}{H}$ | $-\frac{e^2}{c}$ | $\frac{d^2}{c}$ | $+d$ |
| 4 | $x' = +d$ | $-\frac{c}{H}$ | 0 | $-\frac{d}{H}$ | $+\frac{d^2}{c}$ | $\frac{e^2}{c}$ | $+d$ |
| | $y' = -d$ | 0 | $-\frac{c}{H}$ | $+\frac{d}{H}$ | $-\frac{e^2}{c}$ | $-\frac{d^2}{c}$ | $+d$ |

## Discussion

1. The normal equation matrix is sparse. It collapses to two diagonal elements and two $2 \times 2$ matrices. This allows algebraic inversion (which may be used for a direct solution of the orientation in real time applications).
2. The correlation between $x_0$ and $d\phi$ ($y$-rotation), $y_0$ and $d\omega$ ($x$-rotation) is given by

$$\rho_{y_0\omega} = -\rho_{x_0\phi} = \frac{N_{24}}{\sqrt{N_{22} \cdot N_{44}}} = \frac{4\frac{e^2}{H}}{\sqrt{4\frac{c^2}{H^2} \cdot 4\frac{e^4}{c^2}(1 + \frac{d^4}{e^4})}} = \frac{1}{\sqrt{1 + \sin^4 \frac{\alpha}{2}}} \quad (38)$$

   as $d/e = \sin \alpha/2$ (cf. Fig. 7).
3. Taking the square root $\boldsymbol{N}^{1/2}$ of $\boldsymbol{N}$ is trivial for the diagonal elements for $dz_0$ and $d\kappa$ and requires to take the square root $\boldsymbol{T}^{1/2}$ of two $2\times 2$ matrices $\boldsymbol{T}$ which easily can be determined using the eigenvalue decomposition $\boldsymbol{T} = \boldsymbol{D}\boldsymbol{\Lambda}\boldsymbol{D}^T$ yielding

$$\boldsymbol{T}^p = \boldsymbol{D}\boldsymbol{\Lambda}^p\boldsymbol{D}^T \quad (39)$$

for $p = 1/2$ or, as needed in (10) for $p = -1/2$ ($\boldsymbol{D}^T = \boldsymbol{D}^{-1}$ and $\boldsymbol{\Lambda} = \mathrm{Diag}(\lambda_1, \lambda_2)$).

## References

Baarda W.: Statistical Concepts in Geodesy. Netherlands geodetic Commissison, New Series, Vol.2, No. 4., Delft 1967

Baarda W.: A Testing Procedure for Usen in Geodetic Networks. Netherlands geodetic Commissison, New Series, Vol.5, No. 1., Delft 1968

Baarda W.: S-Transformations and Criterion Matrices. Netherlands Geod. Commission New Series, Vol. 5, No. 1, Delft, 1973

Bolles R. C., Fischler M. A.: Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. Comm. of ACM, Vol. 24, No. 6, 1981, 381–395

Chen C.-H., Mulgaonkar P. G.: Robust Vision-Programs Based on Statistical Feature Measures. Int. Workshop on Robust Computer Vision, Proc., Seattle, October 1990, 39–56

Cook R. D., Weisberg S.: Residuals and Influence in Regression. Chapmann & Hall, 1982

Deriche R., Faugeras O.: Tracking Line Segments. Lecture Notes in Computer Science, 427, Springer, 1990, 259–268

Förstner W.: Reliability and Discernability of Extended Gauss-Markov Models. Deutsche Geod. Komm, A98, München, 1983, 79–103

Förstner W.: Reliability Analysis and Parameter Estimation in Linear Models with Application to Mensuration Problems in Computer Vision. CVGIP, 40, 1987, 87–104

Förstner W.: Uncertain Geometric Relationships and their Use for Object Location in Digital Images. Institut für Photogrammetrie, Universität Bonn, Tutorial, 1992

Hampel F. R., Ronchetty E. M., Rousseeuw P. J., Stahel W. A.: Robust Statistics: The Approach Based on Influence Functions. Wiley, NY, 1986

Horaud R., Conio B., Leboulleux O., Lacolle B.: An Analytic Solution for the Perspective 4-Point Problem. CVPR 1989, Proc., San Diego, CA, 1989, 500–507

Huber P. J.: Robust Statistics. Wiley, NY, 1981

Huber P. J.: Between Robustness and Diagnostics. In: Directions in Robust Statistics and Diagnostics, Part I, W. Stahel and S. Weisberg (eds.), Springer Verlag, 1991, 121–130

Roth G., Levine M. D.: Random Sampling for Primitive Extraction. Int. Workshop on Robust Computer Vision, Proc., Seattle, October 1990, 352–366

Rousseeuw P. J., Leroy A. M.: Robust Regression and Outlier Detection. Wiley, NY, 1987

Schickler W.: Feature Matching for Outer Orientation of Single Images Using 3-D Wireframe Controlpoints. Int. Arch. of Photogrammetry & Remote Sensing, Vol. 29, ISPRS Comm. III, Washington D. C., 1992

Sester M., Förstner W.: Object Location Based on Uncertain Models Informatik Fachberichte 219, Springer, 1989, 457–464

Stockman G.: Object Recognition and Localization via Pose Clustering. CVGIP, 40, 1987, 361–387

This article was processed using the LaTeX macro package with LMAMULT style