

minimal representation of homogeneous coordinate, including tangent space, covariance, making M-test available, improve numeric stability  
However, the application for Vanishing point, and 3D line fitting is not very clear, needs to read more papers to get it,

# Minimal Representations for Uncertainty and Estimation in Projective Spaces

Wolfgang Förstner

University Bonn, Department of Photogrammetry  
Nussallee 15, D-53115 Bonn, wf@ipb.uni-bonn.de

**Abstract.** Estimation using homogeneous entities has to cope with obstacles such as **singularities of covariance matrices** and **redundant parameterizations** which do not allow an immediate definition of maximum likelihood estimation and lead to estimation problems with more parameters than necessary. The paper proposes a representation of the **uncertainty of all types of geometric entities and estimation procedures for geometric entities and transformations** which (1) only require the minimum number of parameters, (2) are free of singularities, (3) allow for a consistent update within an iterative procedure, (4) enable to exploit the simplicity of homogeneous coordinates to represent geometric constraints and (5) allow to handle geometric entities which are at infinity or at least very far, **avoiding the usage of concepts like the inverse depth**. Such representations are already available for transformations such as rotations, motions (Rosenhahn 2002), homographies (Begelfor 2005), or the projective correlation with fundamental matrix (Bartoli 2004) all being elements of some Lie group. The uncertainty is represented in the **tangent space of the manifold**, namely the corresponding Lie algebra. However, to our knowledge no such representations are developed for the basic geometric entities such as points, lines and planes, as in addition to use the tangent space of the manifolds we need transformation of the entities such that they stay on their specific manifold during the estimation process. We develop the concept, discuss its usefulness for bundle adjustment and demonstrate (a) its superiority compared to more simple methods for vanishing point estimation, (b) its rigour when estimating 3D lines from 3D points and (c) its applicability for **determining 3D lines from observed image line segments in a multi view setup**.

*Motivation.* Estimation of entities in projective spaces, such as points or transformations, has to cope with the *scale ambiguity* of these entities, resulting from the redundancy of the projective representations, and with the definition of *proper metrics* which on one hand *reflect the uncertainty* of the entities and on the other hand lead to estimates which are *invariant* to possible gauge transformation, thus changes of the reference system. The paper shows how to consistently perform Maximum likelihood estimation for an arbitrary number of geometric entities in projective spaces including elements at infinity under realistic assumptions without a need to **impose constraints resulting from the redundant representations**.

The *scale ambiguity* of homogeneous entities results from the redundant representation, where two elements, say 2D points,  $\chi(\mathbf{x})$  and  $\chi(\mathbf{y})$  are identical, in case their representations with homogeneous coordinates, here with  $\mathbf{x}$  and  $\mathbf{y}$ , are proportional. This ambiguity regularly is avoided by proper normalization of the homogeneous entities. Mostly one applies either Euclidean normalization, say  $\mathbf{x}^e = \mathbf{x}/x_3$  (cf. [13]), then accepting that no elements at infinity can be represented, or spherical normalization, say  $\mathbf{x}^s = \mathbf{x}/|\mathbf{x}|$ , then accepting that the parameters to be estimated sit on a non-linear manifold, here the unit sphere  $S^2$ , cf. [7, 11]. The sign ambiguity usually does not cause difficulty, as the homogeneous constraints used for reasoning are independent on the chosen sign. The singularity of constrained observations also has also been pointed out in [5].

The *uncertainty* of an observed geometric entity in many practical cases, can be represented sufficiently well by a Gaussian distribution, say  $N(\mu_x, \Sigma_{xx})$ . The distribution of derived entities,  $\mathbf{y} = \mathbf{f}(\mathbf{x})$ , resulting from a non-linear transformation can also be approximated by a Gaussian distribution, using Taylor expansion at the mean  $\mu_x$  and omitting higher order terms. The degree of approximation depends on the relative accuracy and has been shown to be negligible in most cases, cf. [8, p. 55].

The invariance of estimates w.r.t. the choice of the coordinate system of the estimated entities usually is achieved, by minimizing a function in the Euclidean space of observations, in the context of bundle adjustment being the reprojection error, leading to the optimization function  $\Omega = \sum_i (\underline{\mathbf{x}}_i - \hat{\underline{\mathbf{x}}}_i)^\top \Sigma_{x_i x_i}^{-1} (\underline{\mathbf{x}}_i - \hat{\underline{\mathbf{x}}}_i)$ . This at the same time is the Mahalanobis distance between the observed and estimated entities and can be used to evaluate whether the model fits to the data.

This situation becomes difficult, in case one wants to handle elements at infinity and therefore wants to use spherically normalized homogeneous vectors, or at least normalized direction vectors when using omnidirectional cameras, as their covariance matrices are singular. The rank deficiency is at least one, due to the homogeneity. In case further constraints need to be taken into account, as the Plücker constraint for 3D lines or the singularity constraint for fundamental matrices, the rank deficiency increases with the number of these constraints. Therefore, in case we want to use these normalized vectors or matrices as observed quantities, already the formulation of the optimization function based on homogeneous entities is not possible and requires a careful discussion about estimable quantities, cf. [17]. Also the redundant representation requires additional constraints, which lead to Lagrangian parameters in the estimation process. As an example, one would need four parameters to estimate a 2D point, three for the homogeneous coordinates and one Lagrangian for the constraint, two parameters more than the degrees of freedom.

It remains open, how to arrive at a minimal representation for the uncertainty and at the estimation of all types of geometric entities in projective spaces which are free of singularities and allow to handle entities at infinity.

*Related work.* This problem has been addressed successfully for geometric transformations. Common to these approaches is the observation that all types of

transformations form differentiable groups called Lie groups. Starting from an approximate transformation, the estimation can exploit the corresponding Lie algebra, being the tangent space at the unit element of the manifold of the group. Take as an example the group  $SO(3)$  of rotations: Starting from an approximate rotation  $R^a$  a close-by rotation  $R$  can be represented by  $R = R(\Delta\mathbf{R})R^a$ , with a small rotation with rotation matrix  $R(\Delta\mathbf{R})$  depending on the small rotation vector  $\Delta\mathbf{R}$ . This small rotation vector can be estimated in the approximate model  $R \approx (I_3 + S(\Delta\mathbf{R}))R^a$ , where the components of the rotation vector  $\Delta\mathbf{R}$  appear linearly, building the three dimensional space of the Lie algebra  $so(3)$  corresponding to the Lie group  $SO(3)$ . Here  $S(\cdot)$  is the skew symmetric matrix inducing the crossproduct. The main difference to a Taylor approximation, which is additive, say  $R \approx R^a + \Delta\mathbf{R}$ , is the multiplicative correction in (1), which guarantees that the corrected matrix is a proper rotation matrix, based on the exponential representation  $R(\Delta\mathbf{R}) = \exp(S(\Delta\mathbf{R}))$  of a rotation using skew symmetric matrices.

This concept can be found in all proposals for a minimal representation for transformations: Based on the work of Bregler et al. [6], Rosenhahn et al. [19] used the exponential map for modelling spatial Euclidean motions, the special Euclidean group  $SE(3)$  being composed of rotations  $SO(3)$  and translations in  $\mathbb{R}^3$ . Bartoli and Sturm [2] used the idea to estimate the fundamental matrix with a minimal representation  $F = R_1 \text{Diag}(\exp(\lambda), \exp(-\lambda), 0) R_2$ , twice using the rotation group and once the multiplication group  $\mathbb{R}^+$ . Begelfor and Werman [4] showed how to estimate a general 2D homography with a minimal representation statistically rigorous, namely using the special linear group  $SL(3)$  of  $3 \times 3$ -matrices with determinant 1, and its Lie algebra  $sl(3)$  consisting of all matrices with trace zero, building an eight dimensional vector space, correctly reflecting the correct number of degrees of freedom.

To our knowledge the only attempt to use minimal representations for geometric entities other than transformations have been given by Sturm [20] and Åström [1]. Sturm suggested a minimal representation of conics, namely  $C = RDR^T$ , with  $D = \text{Diag}(a, b, c)$ . Using a corresponding class of homographies  $H = Q \text{Diag}(d, e, f)$ ,  $Q$  being a rotation matrix, which map any conic into the unit circle  $\text{Diag}(1, 1, -1)$ , he determines updates for the conic in this reduced representation and at the end undoes the mapping. Another way to achieve a minimal representation for homogeneous entities is given by Åström [1] in the context of structure from motion. He proposes to use the Cholesky decomposition of the pseudo inverse of the covariance matrix of the spherically normalized homogeneous 2D point coordinates  $\mathbf{x}$  to arrive at a whitened and reduced observation. For 3D lines he uses a special double point representation with minimal parameters. He provides no method to update the approximate values  $\mathbf{x}^a$  guaranteeing the estimate  $\hat{\mathbf{x}} \in S^2$ . His 3D line representation is also not linked to projective Plücker representation, and he cannot estimate elements at infinity.

Our proposal is similar in flavour to the idea of Sturm and the method of Åström to represent uncertain 2D points. However, it is simpler to handle, as it directly works on the manifold of the homogeneous entities.

*Notation.* We name objects with calligraphic letters, say a point  $\chi$ , in order to be able to change representation, e. g. using Euclidean coordinates denoted with a slanted letter  $\boldsymbol{x}$  or homogeneous coordinates with an upright letter  $\mathbf{x}$ . Matrices are denoted with sans serif capital letters, say  $R$ , or in case of homogeneous matrices  $H$ . This transfers to the indices of covariance matrices, the covariance matrix  $\Sigma_{xx}$  for a Euclidean vector and  $\Sigma_{\mathbf{xx}}$  for a homogeneous vector. The operator  $N(\cdot)$  normalizes a vector to unit length. We adopt the Matlab syntax to denote the stack of two vectors or matrices, e. g.  $\mathbf{z} = [\mathbf{x}; \mathbf{y}] = [\mathbf{x}^T, \mathbf{y}^T]^T$ . Stochastic variables are underscored, e. g.  $\underline{\mathbf{x}}$ .

## 1 Minimal Representation of Uncertainty

The natural space of homogeneous entities are the unit spheres  $S^n$ , possibly constrained to a subspace. Spherically normalized homogeneous coordinates of 2D points ( $\mathbf{x}^s$ ) and lines ( $\mathbf{l}^s$ ), and of 3D points ( $\mathbf{X}^s$ ) and planes ( $\mathbf{A}^s$ ) live on  $S^2$  and  $S^3$  resp., while 3D lines, represented by Plücker coordinates ( $\mathbf{L}^s$ ), live on the Klein quadric  $Q$ . The points on  $S^3$  build the Lie group of unit quaternions, which could be used for an incremental update  $\mathbf{q} = \Delta \mathbf{q} \mathbf{q}^a$  of spherically normalized 3D point or plane vectors. However, there exist no Lie groups for points on the 2-sphere or on the Klein quadric, see [18]. Therefore we need to develop an update scheme, which guarantees the updates to lie on the manifold, without relying on some group concept.

We will develop the concept for unit vectors on the  $S^2$ , representing 2D points and lines, and generalize it to the other geometric entities.

### 1.1 Minimal representation for uncertain points in 2D and 3D

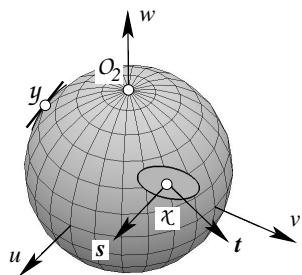
Let an uncertain 2D point  $\chi$  be represented with its mean, the 2-vector  $\boldsymbol{\mu}_x$  and its  $2 \times 2$ -covariance matrix  $\Sigma_{xx}$ . It can be visualized by the standard ellipse  $(\mathbf{x} - \boldsymbol{\mu}_x)^T \Sigma_{xx}^{-1} (\mathbf{x} - \boldsymbol{\mu}_x) = 1$ . Spherically normalizing the homogeneous vector  $\mathbf{x} = [\mathbf{x}; 1] = [u, v, w]^T$  yields

$$\mathbf{x}^s = \frac{\mathbf{x}}{|\mathbf{x}|}, \quad \Sigma_{x^s x^s} = J \Sigma_{xx} J^T \quad \text{with} \quad \Sigma_{xx} = \begin{bmatrix} \Sigma_{xx} & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix}, \quad J = \frac{1}{|\mathbf{x}|} (I_3 - \mathbf{x}^s \mathbf{x}^{sT}) \quad (1)$$

with  $\text{rank}(\Sigma_{xx}) = 2$  and  $\text{null}(\Sigma_{xx}) = \mathbf{x}^s$ . Taking the smallest eigenvalue to be an infinitely small positive number, one sees that the standard error ellipsoid is flat and lies in the tangent space of  $\mathbf{x}$  at  $S^2$ . In the following we assume all point vectors  $\mathbf{x}$  to be spherically normalized and omit the superscript  $s$  for simplicity of notation.

We now want to choose a coordinate system  $J_x(\mathbf{x}) = [\mathbf{s}, \mathbf{t}]$  in the tangent space  $\perp \mathbf{x}$  and represent the uncertainty by a  $2 \times 2$ -matrix in that coordinate system. This is easily achieved by using

$$J_x(\mathbf{x}) = \text{null}(\mathbf{x}^T), \quad (2)$$



**Fig. 1.** Minimal representation for an uncertain point  $\chi(\mathbf{x})$  on the unit sphere  $S^2$  representing the projective plane  $\mathbb{P}^2$  by a flat ellipsoid in the tangent plane at  $\chi$ . The uncertainty has only two degrees of freedom in the tangent space spanned by two basis vectors  $s$  and  $t$  of the tangent space, being the null space of  $\mathbf{x}^\top$ . The uncertainty should not be too large, such that the deviation of the distribution on the sphere and on the tangent do not differ too much, as at point  $y$ .

assuming it to be an orthonormal matrix fulfilling  $J_x^\top(\mathbf{x}) J_x(\mathbf{x}) = I_2$ , see fig. 1. We define a stochastic 2-vector  $\underline{\mathbf{x}}_r \sim M(\mathbf{0}, \Sigma_{x_r x_r})$  with mean  $\mathbf{0}$  and covariance  $\Sigma_{x_r x_r}$  in the tangent space at  $\chi$ . In order to arrive at a spherically normalized random vector  $\underline{\mathbf{x}}$  with mean  $\mu_x$  we need to spherically normalize the vector

$$\underline{\mathbf{x}}^t = \mu_x + J_x(\mu_x) \underline{\mathbf{x}}_r \quad (3)$$

in the tangent space and obtain

$$\underline{\mathbf{x}}(\mu_x, \underline{\mathbf{x}}_r) = N(\mu_x + J_x(\mu_x) \underline{\mathbf{x}}_r), \quad J_x(\mu_x) = \frac{\partial \mathbf{x}}{\partial \mathbf{x}_r} \Big|_{x=\mu_x} \quad (4)$$

We thus can identify  $J_x(\mu_x)$  with the Jacobian of this transformation. The inverse transformation can be achieved using the pseudo inverse of  $J_x(\mathbf{x})$  which due to the construction is the transpose,  $J^+(\mathbf{x}) = J_x^\top(\mathbf{x})$ . This leads to the reduction of the homogeneous vector to its reduced counterpart

$$\underline{\mathbf{x}}_r = J_x^\top(\mu_x) \underline{\mathbf{x}}. \quad (5)$$

As  $J_x^\top(\mu_x) \mu_x = \mathbf{0}$  the mean of  $\underline{\mathbf{x}}_r$  is the zero vector,  $\mu_{x_r} = \mathbf{0}$ .

This allows to establish the one to one correspondence between the *reduced covariance matrix*  $\Sigma_{x_r x_r}$  and the covariance matrix  $\Sigma_{xx}$  of  $\underline{\mathbf{x}}$ :

$$\Sigma_{xx} = J_x(\mu_x) \Sigma_{x_r x_r} J_x^\top(\mu_x), \quad \Sigma_{x_r x_r} = J_x(\mu_x)^\top \Sigma_{xx} J_x(\mu_x). \quad (6)$$

We use (5) to derive *reduced* observations and parameters and after estimating corrections  $\widehat{\Delta \mathbf{x}}_r$  then apply (4) to find corrected estimates  $\widehat{\mathbf{x}} = \widehat{\mathbf{x}}(\mathbf{x}^a, \widehat{\Delta \mathbf{x}}_r)$ .

A similar reasoning leads to the representation of 3D points. Again, the Jacobian  $J_X$  is the null space of  $\mathbf{X}^\top$  and spans the 3-dimensional tangent space of  $S^3$  at  $\mathbf{X}$ . The relations between the singular  $4 \times 4$ -covariance matrix of the spherically normalized vector  $\mathbf{X}$  and the reduced  $3 \times 3$ -covariance matrix  $\Sigma_{X_r X_r}$  are equivalent to (6).

Homogeneous 3-vectors  $\mathbf{l}$  representing 2D lines and homogeneous 4-vectors  $\mathbf{A}$  representing planes can be handled in the same way.

## 1.2 Minimal representation for 3D lines

We now generalize the concept for 3D lines. Lines  $\mathcal{L}$  in 3D are represented by their Plücker coordinates  $\mathbf{L} = [\mathbf{L}_h; \mathbf{L}_0] = [\mathbf{Y} - \mathbf{X}, \mathbf{X} \times \mathbf{Y}]$  in case they are

derived by joining two points  $\mathcal{X}(\mathbf{X})$  and  $\mathcal{Y}(\mathbf{Y})$ . Line vectors need to fulfill the quadratic Plücker constraint  $\mathbf{L}_h^T \mathbf{L}_0 = 0$  and span the Klein quadric  $Q$  consisting of all homogeneous 6-vectors fulfilling the Plücker constraint. The dual line  $\bar{\mathbf{L}}(\bar{\mathbf{L}})$  has Plücker coordinates  $\bar{\mathbf{L}} = [\mathbf{L}_0; \mathbf{L}_h]$ , exchanging its first and second 3-vector. We also will use the Plücker matrix  $\Gamma(\mathbf{L}) = \mathbf{XY}^T - \mathbf{YX}^T$  of a line and will assume 3D line vectors  $\mathbf{L}$  to be spherically normalized. As in addition a 6-vector needs to fulfill the Plücker constraint in order to represent a 3D line, the space of 3D lines is four dimensional.

The transfer of the minimal representation of points to 3D-lines requires some care. The four dimensional tangent space is perpendicular to  $\mathbf{L}$ , as  $\mathbf{L}^T \mathbf{L} - 1 = 0$  holds and perpendicular to  $\bar{\mathbf{L}}$ , as  $\bar{\mathbf{L}}^T \mathbf{L} = 0$  holds. Therefore the tangent space is given by the four columns of the  $6 \times 4$  matrix

$$J_L(\mathbf{L}) = \text{null} \left( [\mathbf{L}, \bar{\mathbf{L}}]^T \right) \quad (7)$$

again assuming this matrix to be orthonormal. Therefore for random perturbations  $\underline{\mathbf{L}}_r$  we have the general 6-vector

$$\underline{\mathbf{L}}^t(\boldsymbol{\mu}_L, \underline{\mathbf{L}}_r) = \boldsymbol{\mu}_L + J_L(\boldsymbol{\mu}_L) \underline{\mathbf{L}}_r \quad (8)$$

in the tangent space.

In order to arrive at a random 6-vector, which is both spherically normalized and fulfills the Plücker constraint also for finite random perturbations we need to normalize  $\mathbf{L}^t = [\mathbf{L}_h^t, \mathbf{L}_0^t]$  accordingly. The two 3-vectors  $\mathbf{L}_h^t$  and  $\mathbf{L}_0^t$  in general are not orthogonal. Following the idea of Bartoli [3] we therefore rotate these vectors in their common plane such that they become orthogonal. We use a simplified modification, as the normalization within an iteration sequence will have decreasing effect. We use linear interpolation of the directions  $\mathbf{D}_h = \mathbf{N}(L_h^t)$  and  $\mathbf{D}_0 = \mathbf{N}(L_0^t)$ . With the distance  $d = |\mathbf{D}_h - \mathbf{D}_0|$  and the shortest distance  $r = \sqrt{1 - d^2/4}$  of the origin to the line joining  $\mathbf{D}_h$  and  $\mathbf{D}_0$  we have

$$\mathbf{M}_{h,0} = ((d \pm 2r) \mathbf{D}_h + (d \mp 2r) \mathbf{D}_0) |\mathbf{L}_h^t| / (2d) \quad (9)$$

The 6-vector  $\mathbf{M} = [\mathbf{M}_h; \mathbf{M}_0]$  now fulfills the Plücker constraint but needs to be spherically normalized. This finally leads to the normalized stochastic 3D line coordinates

$$\underline{\mathbf{L}} = \mathbf{N}(\underline{\mathbf{L}}^t(\boldsymbol{\mu}_L, \underline{\mathbf{L}}_r)) \doteq \underline{\mathbf{M}} / |\underline{\mathbf{M}}| \quad (10)$$

which guarantees  $\underline{\mathbf{L}}$  to sit on the Klein quadric, thus to fulfill the Plücker constraint.

zed, omitting the superscript  $e$  for clarity.

The inverse relation to (10) is

A little confuse: it seems assume  $\mathbf{L} \sim \mathbf{L}^t$ , when compute covariance

$$\underline{\mathbf{L}}_r = J_L^T(\boldsymbol{\mu}_L) \underline{\mathbf{L}} \quad (11)$$

as  $J_L(\boldsymbol{\mu}_L)$  is an orthonormal matrix. The relation between the covariances of  $\mathbf{L}$  and  $\underline{\mathbf{L}}_r$  therefore are

$$\Sigma_{LL} = J_L(\boldsymbol{\mu}_L) \Sigma_{L_r L_r} J_L^T(\boldsymbol{\mu}_L), \quad \Sigma_{L_r L_r} = J_L(\boldsymbol{\mu}_L)^+ \Sigma_{LL} J_L^{+T}(\boldsymbol{\mu}_L). \quad (12)$$

Conics and quadrics can be handled in the same manner. A conic  $C = [a, b, c; b, d, e; c, d, e]$  can be represented by the vector  $\mathbf{c} = [a, b, c, d, e, f]^\top$ , normalized to 1, and – due to the homogeneity of the representation – is a point in  $\mathbb{P}^5$  with the required Jacobian being the nullspace of  $\mathbf{c}^\top$ . In contrast to the approach of Sturm [20], this setup would allow for singular conics.

Using the minimal representations introduced in the last section, we are able to perform ML-estimation for all entities. We restrict the model to contain constraints between observed and unknown entities only, not between the parameters only; a generalization is easily possible. We start with the model where the observations have regular covariance matrices and then reduce the model, such that also observations with singular covariance matrices can be handled. Finally, we show how to arrive at a Mahalanobis distance for uncertain geometric entities, where we need the inverse of the covariance matrix.

*The optimization problem.* We want to solve the following optimization problem

$$\text{minimize } \Omega(\mathbf{v}) = \mathbf{v}^\top \Sigma_{ll}^{-1} \mathbf{v} \quad \text{subject to } \mathbf{g}(l + \mathbf{v}, \mathbf{x}) = \mathbf{0} \quad (13)$$

where the  $N$  observations  $\mathbf{l}$ , their  $N \times N$  covariance matrix  $\Sigma_{ll}$  and the  $G$  constraint functions  $\mathbf{g}$  are given, and the  $N$  residuals  $\mathbf{v}$  and the  $U$  parameters  $\mathbf{x}$  are unknown. The number  $G$  of constraints needs to be larger than the number of parameters  $U$ . Also it is assumed the constraints are functionally independent. The solution yields the maximum likelihood estimates, namely the fitted observations  $\hat{\mathbf{x}}$  and parameters  $\hat{\mathbf{l}}$ , under the assumption that the observations are normally distributed with covariance matrix  $\Sigma_{ll} = D(\mathbf{l}) = D(\mathbf{v})$ , and the true observations  $\tilde{\mathbf{l}}$  fullfill the constraints given the true parameters  $\tilde{\mathbf{x}}$ .

**Example: Bundle adjustment.** Bundle adjustment is based on the projection relation  $\mathbf{x}'_{ij} = \lambda_{ij} \mathbf{P}_j \mathbf{X}_i$  between the scene points  $\mathbf{X}_i$ , the projection matrices  $\mathbf{P}_j$  and the image points  $\mathbf{x}_{ij}$  of point  $\mathbf{X}_i$  observed in camera  $j$ . The classical approach eliminates the individual scale factors  $\lambda_{ij}$  by using Euclidean coordinates for the image points. Also the scene points are represented by their Euclidean coordinates. This does not allow for scene or image points at infinity. This may occur when using omnidirectional cameras, where a representation of the image points in a projection plane is not possible for all points or in case scene points are very far compared to the length of the motion path of a camera, e.g. at the horizon. Rewriting the model as  $\mathbf{x}_{ij} \times \mathbf{P}_j \mathbf{X}_i = \mathbf{0}$  eliminates the scale factor without constraining the image points to be not at infinity. Taking the homogeneous coordinates of all image points as observations and the parameters of all cameras and the homogeneous coordinates of all scene points as unknown parameters shows this model to have the structure of (13). The singularity of the covariance matrix of the spherically normalized image points and the necessity to represent the scene points also with spherically normalized homogeneous vectors, requires to use the corresponding reduced coordinates.  $\diamond$

For solving the generally non-linear problem, we assume approximate values  $\mathbf{x}^a$  and  $\hat{\mathbf{l}}^a$  for the fitted parameters and observations to be available. We thus search for corrections  $\Delta\mathbf{l}$  and  $\Delta\mathbf{x}$  for the fitted observations and parameters using

$\hat{\mathbf{l}} = \mathbf{l} + \hat{\mathbf{v}} = \hat{\mathbf{l}}^a + \widehat{\Delta\mathbf{l}}$  and  $\hat{\mathbf{x}} = \hat{\mathbf{x}}^a + \widehat{\Delta\mathbf{x}}$ . With these assumptions we can rephrase the optimization problem: minimize  $\Omega(\widehat{\Delta\mathbf{l}}) = (\hat{\mathbf{l}}^a - \mathbf{l} + \widehat{\Delta\mathbf{l}})^\top \Sigma_{ll}^{-1} (\hat{\mathbf{l}}^a - \mathbf{l} + \widehat{\Delta\mathbf{l}})$  subject to  $\mathbf{g}(\hat{\mathbf{l}}^a + \widehat{\Delta\mathbf{l}}, \hat{\mathbf{x}}^a + \widehat{\Delta\mathbf{x}}) = \mathbf{0}$ . The approximate values are iteratively improved by finding best estimates for  $\widehat{\Delta\mathbf{l}}$  and  $\widehat{\Delta\mathbf{x}}$  using the linearized constraints

$$\mathbf{g}(\hat{\mathbf{l}}^a, \hat{\mathbf{x}}^a) + A\widehat{\Delta\mathbf{x}} + B^\top \widehat{\Delta\mathbf{l}} = \mathbf{0} \quad (14)$$

with the corresponding Jacobians  $A$  and  $B$  of  $\mathbf{g}$  to be evaluated at the approximate values.

*Reducing the model.* We now want to transform the model in order to allow for observations with singular covariances. For simplicity we assume the vectors  $\mathbf{l}$  and  $\mathbf{x}$  of all observations and unknown parameters can be partitioned into  $I$  and  $J$  individual and mutually uncorrelated observational vectors  $\mathbf{l}_i, i = 1, \dots, I$  and parameter vectors  $\mathbf{x}_j, j = 1, \dots, J$ , referring to points, lines, planes or transformations.

We first introduce the reduced observations  $\mathbf{l}_{ri}$ , the reduced corrections of the observations  $\widehat{\Delta\mathbf{l}}_{ri}$  and the reduced corrections  $\widehat{\Delta\mathbf{x}}_{rj}$ :

$$\mathbf{l}_{ri} = J_{li}^\top (\hat{\mathbf{l}}^a, \hat{\mathbf{x}}^a) \mathbf{l}_i, \quad \widehat{\Delta\mathbf{l}}_{ri} = J_{li}^\top (\hat{\mathbf{l}}^a, \hat{\mathbf{x}}^a) \widehat{\Delta\mathbf{l}}_i, \quad \widehat{\Delta\mathbf{x}}_{rj} = J_{xj}^\top (\hat{\mathbf{l}}^a, \hat{\mathbf{x}}^a) \widehat{\Delta\mathbf{x}}_j \quad (15)$$

where each Jacobian refers to the type of the entity it is applied to. The reduced approximate values are zero, as they are used to define the reduction, e. g. from (5) we conclude  $\mathbf{x}_r^a = J_x^\top (\mathbf{x}^a) \mathbf{x}^a = \mathbf{0}$ . We collect the Jacobians in two block diagonal matrices  $J_l^\top = \{J_{li}^\top(\hat{\mathbf{l}}^a, \hat{\mathbf{x}}^a)\}$  and  $J_x^\top = \{J_{xj}^\top(\hat{\mathbf{l}}^a, \hat{\mathbf{x}}^a)\}$  in order to arrive at the reduced observations  $\mathbf{l}_r = J_l^\top \mathbf{l}$ , the corrections for the reduced observations  $\Delta\mathbf{l}_r = J_l^\top \Delta\mathbf{l}$  and parameters  $\Delta\mathbf{x}_r = J_x^\top \Delta\mathbf{x}$ .

Second we need to reduce the covariance matrices  $\Sigma_{l_i l_i}$ . This requires some care: As a covariance matrix is the mean squared deviation from the mean, we need to refer to the best estimate of the mean when using it. In our context the best estimate for the mean at the current iteration is the approximate value  $\hat{\mathbf{l}}_i^a$ . Therefore we need to apply two steps: (1) transfer the given covariance matrix, referring to  $\mathbf{l}_i$ , such that it refers to  $\hat{\mathbf{l}}_i^a$  and (2) reduce the covariance matrix to the minimal representation  $\hat{\mathbf{l}}_{ri}$ . As an example, let the observations be 2D lines with spherically normalized homogeneous vectors  $\mathbf{l}_i$ . Then the reduction is achieved by:  $\Sigma_{l_r l_i}^a = J_i^a \Sigma_{l_i l_i} J_i^{a\top}$  with  $J_i^a = J_x^\top(\mathbf{l}_i^a) R(\mathbf{l}_i, \hat{\mathbf{l}}_i^a)$ , namely by first applying a minimal rotation from  $\mathbf{l}_i$  to  $\hat{\mathbf{l}}_i^a$  (see [16, eq. (2.183)], second reducing the covariance matrix following (6)). Observe, we use the same Jacobian  $J_x$  as for points, exploiting the duality of 2D points and 2D lines. The superscript  $a$  in  $\Sigma_{l_r l_i}^a$  indicates the covariance to depend on the approximated values.

The reduced constraints now read as

$$\mathbf{g}(\hat{\mathbf{l}}^a, \hat{\mathbf{x}}^a) + A_r \widehat{\Delta\mathbf{x}}_r + B_r^\top \widehat{\Delta\mathbf{l}}_r = \mathbf{0} \quad \text{with} \quad A_r = AJ_x^\top \quad B_r^\top = B^\top J_l^\top \quad (16)$$

Now we need to minimize the weighted sum of the squared reduced residuals  $\hat{\mathbf{v}}_r = \hat{\mathbf{l}}_r^a - \mathbf{l}_r + \widehat{\Delta\mathbf{l}}_r = -\mathbf{l}_r + \widehat{\Delta\mathbf{l}}_r$  thus we need to minimize  $\Omega(\widehat{\Delta\mathbf{l}}_r) = (-\mathbf{l}_r + \widehat{\Delta\mathbf{l}}_r)^\top (\Sigma_{l_r l_r}^a)^{-1} (-\mathbf{l}_r + \widehat{\Delta\mathbf{l}}_r)$  subject to the reduced constraints in (16).

The parameters of the linearized model are obtained from (cf. [16, Tab. 2.3])

How this solution obtained?  $\Sigma_{\hat{x}_r \hat{x}_r} = (\mathbf{A}_r^\top (\mathbf{B}_r^\top \Sigma_{l_r l_r}^a \mathbf{B}_r)^{-1} \mathbf{A}_r)^{-1}$  (17)

$$\widehat{\Delta \mathbf{x}}_r = \Sigma_{\hat{x}_r \hat{x}_r} \mathbf{A}_r^\top (\mathbf{B}_r^\top \Sigma_{l_r l_r}^a \mathbf{B}_r)^{-1} \mathbf{w}_g \quad (18)$$

$$\widehat{\Delta l}_r = \Sigma_{l_r l_r}^a \mathbf{B}_r (\mathbf{B}_r^\top \Sigma_{l_r l_r}^a \mathbf{B}_r)^{-1} (\mathbf{w}_g - \mathbf{A}_r \widehat{\Delta \mathbf{x}}_r) - \widehat{\mathbf{v}}^a \quad (19)$$

using  $\mathbf{w}_g = -\mathbf{g}(\widehat{l}^a, \widehat{\mathbf{x}}^a) + \mathbf{B}_r^\top \widehat{\mathbf{v}}_r$ . It contains the theoretical covariance matrix  $\Sigma_{\hat{x}_r \hat{x}_r}$  in (17), at the same time being the Cramer-Rao-bound. The weighted sum of residuals  $\Omega$  is  $\chi^2$ -distributed with  $G - U$  degrees of freedom, in case the observations fulfill the constraints and the observations are normally distributed with covariance matrix  $\Sigma_{ll}$  and can be used to test the validity of the model. The corrections  $\widehat{\Delta \mathbf{x}}_r$  and  $\widehat{\Delta l}_r$  are used to update the approximate values for the parameters and the fitted observations using the corresponding non-linear transformations, e. g. for an unknown 3D line one uses (10).

**Example: Mahalanobis distance of two 3D lines.** As an example for such an estimation we want to derive the mean of two 3D-lines  $\mathcal{L}_i(\mathbf{L}_i, \Sigma_{L_i L_i})$ ,  $i = 1, 2$  and derive the Mahalanobis distance  $d(\mathbf{L}_1, \mathbf{L}_2)$  of the 3D lines. The model reads  $\mathbf{g}(\widehat{\mathbf{L}}_1, \widehat{\mathbf{L}}_2) = \widehat{\mathbf{L}}_2 - \widehat{\mathbf{L}}_1 = \mathbf{0}$ , which, exploits the fact that the line coordinates are normalized. First we need to choose an appropriate approximate line  $\mathbf{L}^a$ , which in case the two lines are not too different can be one of the two. Then we reduce the two lines using the Jacobian  $J_L(\mathbf{L}^a)$ , being the same for both lines  $\mathcal{L}_i$ , and obtain  $\mathbf{L}_{ri} = J_L^\top(\mathbf{L}^a) \mathbf{L}_i$  and the reduced covariance matrices  $\Sigma_{L_{ri}, L_{ri}}^a = J_i^\top \Sigma_{L_i L_i} J_i$  with  $J_i^\top = J_L^\top(\mathbf{L}^a) R(\mathbf{L}_i, \mathbf{L}^a)$  using the minimal rotation from  $\mathbf{L}_i$  to  $\mathbf{L}^a$ . As there are no parameters to be estimated, the solution becomes simple. With the Jacobian  $\mathbf{B}^\top = [-I_6, I_6]$  and using  $\widehat{\Delta \mathbf{x}}_r = \mathbf{0}$  in (19) the reduced residuals are  $\widehat{\mathbf{v}}_{ri} = \pm \Sigma_{L_{ri}, L_{ri}}^a (\Sigma_{L_{r1}, L_{r1}}^a + \Sigma_{L_{r2}, L_{r2}}^a)^{-1} (-(\mathbf{L}_{r2} - \mathbf{L}_{r1}))$ . The weighted sum  $\Omega$  of the residuals is the Mahalanobis distance and therefore can be determined from

$$d^2(\mathcal{L}_1, \mathcal{L}_2) = (\mathbf{L}_{r2} - \mathbf{L}_{r1})^\top (\Sigma_{L_{r2}, L_{r2}}^a + \Sigma_{L_{r1}, L_{r1}}^a)^{-1} (\mathbf{L}_{r2} - \mathbf{L}_{r1}), \quad (20)$$

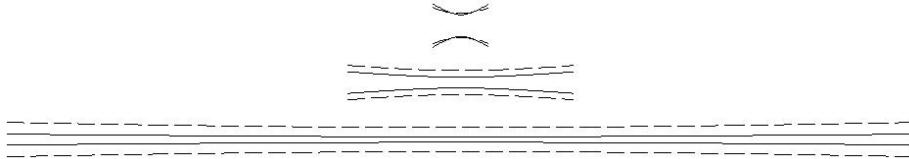
as the reduced covariance matrices in general have full rank. The squared distance  $d^2$  is  $\chi^2$  distributed, which can be used for testing. In case one does not have the covariance matrices of the 3D lines  $\mathcal{L}_i$  but only two points  $\mathcal{X}_i(\mathbf{X}_i)$  and  $\mathcal{Y}_i(\mathbf{Y}_i)$  of a 3D line segment, and a good guess for their uncertainty, say  $\sigma_2 I_3$ , one easily can derive the covariance matrix from  $\mathbf{L}_i = [\mathbf{Y}_i - \mathbf{X}_i; \mathbf{X}_i \times \mathbf{Y}_i]$  by variance propagation. The equations directly transfer to the Mahalanobis distance of two 2D line segments. Thus, no heuristic is required to determine the distance of two line segments.

## 2 Examples

We want to demonstrate the setup with three interrelated problems: (1) demonstrating the superiority of the rigorous estimation compared to a classical one when estimating a vanishing point, (2) fitting a straight line through a set of 3D

points and (3) inferring a 3D line from image line segments observed in several images as these tasks regularly appear in 3D reconstruction of man-made scenes and solved suboptimally, see e.g. [12].

*Estimating a vanishing point.* Examples for vanishing point estimation from line segments using this methodology are found in [9]. Here we want to demonstrate that using a simple optimization criterion can lead to less interpretable results. We compare two models for the random perturbations of line segments:  $\alpha$ ) a model which determines the segment by fitting a line through the edge pixels, which are assumed to be determinable with a standard deviation of  $\sigma_p$ ,  $\beta$ ) the model of Liebowitz [15, Fig. 3.16], who assumes the endpoints to have zero mean isotropic Gaussian distribution, and that each of the endpoint measurements are independent and have the same variance  $\sigma_e^2$ . Fig. 2 demonstrates the difference of the two models. The model  $\alpha$ , using line fitting, results in a significant decrease of the uncertainty for longer line segments, the difference in standard deviation going with the squareroot of the lenght of the line segment. For a simulated case with 10 lines between 14 and 85 pixels, the uncertainty



**Fig. 2.** Error bounds ( $1 \sigma$ ) for line segments. Dashed: following the model of Liebowitz [15], standard deviation standard deviation at end points  $\sigma = 0.15$  [pel] , Solid: from fitting a line through the edge pixels with standard deviation  $\sigma_p = 0.3$  [pel] for segments with 10, 40 and 160 pixels. Deviations 20 times overemphasized. The uncertainty differ by a factor of 2 on the average in this case.

models on an average differ by  $\sqrt[4]{85/14} \approx 1.6$  in standard deviation. Based on 200 repetitions, the empirical scatter of the vanishing point of the method  $\beta$  of Liebowitz is appr. 20 % larger in standard deviation than the method  $\alpha$  using the line fitting accuracy as error model. This is a small gain. However, when statistically testing the line segments whether they belong to the vanishing points, the decision depends on the assumed stochstical model for the line segments: When compared to short segments, the test will be much more sensitive for long segments in case model  $\alpha$  is used, as when model  $\beta$  is used, which appears to be reasonable.

*Fitting a 3D line though points.* We now give an example proving the validity of the approach again using simulated data in order to have full access to the uncertainty of the given data. Given are  $I$  uncertain 3D points  $X_i(\tilde{\mathbf{X}}_i, \Sigma_{X_i X_i}), i = 1, \dots, I$ , whose true values are supposed to sit on a 3D line  $\mathcal{L}$ . The two constraints for the incidence of point  $X_i$  and the line  $\mathcal{L}$  can be written as

$$\mathbf{g}_i(\tilde{\mathbf{X}}_i, \tilde{\mathbf{L}}) = J_1(\tilde{\mathbf{L}}) \tilde{\mathbf{X}}_i = J_2(\tilde{\mathbf{X}}_i) \tilde{\mathbf{L}} = \mathbf{0} \quad (21)$$

with the Jacobians  $J_1(\mathbf{L})$  and  $J_2(\mathbf{X}_i)$  depending on the homogeneous coordinates of the points and the line. The incidence constraint of a point and a line can be

expressed with the Plücker matrix of the dual line:  $\Gamma(\bar{\mathbf{L}})\mathbf{X} = \mathbf{0}$ . As from these four incidence constraints only two are linearly independent, we select those two, where the entries in  $\Gamma(\bar{\mathbf{L}})$  have maximal absolute value, leading to the two constraints  $\mathbf{g}_i(\mathbf{X}_i, \mathbf{L}) = J_1(\mathbf{L})\mathbf{X}_i$  for each point with the  $2 \times 4$  matrix  $J_1(\mathbf{L})$ . The Jacobian  $J_2(\mathbf{X}_i) = \partial \mathbf{g}_i / \partial \mathbf{L}$  then is a  $2 \times 6$ -matrix, cf. [16, sect. 2.3.5.3]. Reduction of these constraints leads to  $\mathbf{g}_i(\mathbf{X}_i, \mathbf{L}) = J_{1r}(\mathbf{L})\mathbf{X}_i = J_{r2}(\mathbf{X}_i)\mathbf{L} = \mathbf{0}$  with the reduced Jacobians  $J_{1r}(\mathbf{L})$  and  $J_{r2}(\mathbf{X}_i)$  having size  $2 \times 3$  and  $2 \times 4$  leading to a  $4 \times 4$  normal equation matrix, the inverse of (17). Observe, the estimation does *not need to incorporate the Plücker constraint*, as this is taken into account after the estimation of  $\mathbf{L}_r$  by the back projection (10).

We want to perform two tests of the setup: (1) whether the estimated variance factor actually is  $F$ -distributed, and (2) whether the theoretical covariance matrix  $\Sigma_{\hat{\mathbf{L}}, \hat{\mathbf{L}}}$  corresponds to the empirical one. For this we define true line parameters  $\tilde{\mathbf{L}}$ , generate  $I$  true points  $\tilde{\mathbf{X}}_i$  on this line, choose individual covariance matrices  $\Sigma_{X_i X_i}$  and perturb the points according to a normal distribution with zero mean and these covariance matrices. We generate  $M$  samples, by repeating the perturbation  $M$  times. We determine initial estimates for the lines using an algebraic minimization, based on the constraints (21), which are linear in  $\mathbf{L}$  and perform the iterative ML-estimation. The iteration is terminated in case the corrections to the parameters are smaller than 0.1 their standard deviation.

We first perform a test where the line passes through the unit cube and contains  $I = 100$  points with, thus  $G = 200$  and  $U = 4$ . The standard deviations of the points vary between 0.0002 and 0.03, thus range to up to 3% relative accuracy referring to the distance to their centroid, which is comparably large. As each weighted sum of squared residuals  $\Omega$  is  $\chi^2$ -distributed with  $G - U = R = 196$  degrees of freedom, the sum  $\overline{\Omega} = \sum_m \omega_m$  of the independent samples  $\Omega_m$  is  $\chi^2_{MR}$  distributed, thus the average variance factor  $\overline{\sigma_0^2} = \sum_m \widehat{\sigma_0}_m^2$  is  $F_{MR, \infty}$  distributed. The histogram of  $M = 1000$  samples  $\widehat{\sigma_0}_m^2$  is shown in fig. 3, left.

Second, we compare the sample covariance matrix  $\widehat{\Sigma}_{\hat{\mathbf{L}}, \hat{\mathbf{L}}}$  with  $\Sigma_{\hat{\mathbf{L}}, \hat{\mathbf{L}}}$ , the Cramer-Rao-bound and a lower bound for the true covariance matrix, using the test for the hypothesis  $H_0 : \widehat{\Sigma}_{\hat{\mathbf{L}}, \hat{\mathbf{L}}} = \Sigma_{\hat{\mathbf{L}}, \hat{\mathbf{L}}}$  from [14, eq. (287.5)]. As both covariance matrices are singular we on one side reduce the theoretical covariance matrix to obtain  $\Sigma_{\hat{\mathbf{L}}_r, \hat{\mathbf{L}}_r}$  and on the other side reduce the estimated line parameters which we use to determine the empirical covariance matrix  $\widehat{\Sigma}_{\hat{\mathbf{L}}_r, \hat{\mathbf{L}}_r}$  from the set  $\{\hat{\mathbf{L}}_{rm}\}$  of the  $M$  estimated 3D lines. The test statistic indicates that the hypothesis of the two covariance matrices being was not rejected.

*Estimating 3D lines from image line segments.* The following example demonstrates the practical use of the proposed method: namely determining 3D lines from image line segments. Fig. 3 shows three images taken with a CANON 450D. The focal length was determined using vanishing points, the principle point was assumed to be the image centre, the images were not corrected for lens distortion. The images then have been mutually oriented using a bundle adjustment program. Straight line segments were automatically detected and a small subset of 12, visible in all three and pointing in the three principle directions were manually

brought into correspondence. From the straight lines  $\ell_{ij}, i = 1, \dots, 12; j = 1, 2, 3$  and the projection matrices  $P_j$  we determine the projection planes  $A_{ij} = P_j^T l_{ij}$ . For determining the ML-estimates of the 12 lines  $L_i$  we need the covariance matrices of the projection planes. They are determined by variance propagation based on the covariance matrices of the image lines  $l_{ij}$  and the covariance matrices of the projection matrices. As we do not have the cross-covariance ma-



**Fig. 3.** Left: Histogram of estimated variance factors  $\widehat{\sigma}_{0m}^{-2}$  determined from  $M = 1000$  samples of a 3D line fitted through 100 points. The 99 % confidence interval for the Fisher distributed random variable  $\widehat{\sigma}_0^{-2}$  is  $[0.80, 1.25]$ . Right: three images with 12 corresponding straight line segments used for the reconstruction of the 3D lines, forming three groups  $[1..4], [5..8], [9..12]$  for three main directions.

trices between any two of the projection matrices, we only use the uncertainty  $\Sigma_{Z_j Z_j}$  of the three projection centres  $Z_j$ . The covariance matrices of the straight line segments are derived from the uncertainty given by the feature extraction as in the first example. The covariance matrix of the projection planes then is determined from  $\Sigma_{A_{ij} A_{ij}} = P_i^T \Sigma_{l_{ii} l_{ii}} P_i + (I_4 \otimes l_{ij}) \Sigma_{p_i p_i} (I_4 \otimes l_{ij}^T)$ .

Now we observe, that determining the intersecting line of three planes is dual to determining the best fitting line through three 3D points. Thus the procedure for fitting a 3D line through a set of 3D points can be used directly to determine the ML-estimate of the best fitting line through three planes, except for the fact, that the result of the estimation yields the dual line. First, the square roots of the estimated variance factors  $\widehat{\sigma}_0^2 = \Omega/(G - U)$  range between 0.03 and 3.2. As the number of degrees of freedom is  $G - U = 2I - 4 = 2 \cdot 3 - 4 = 2$  in this case is very low, such a spread is to be expected. The mean value for the variance factor is 1.1, which confirms the model to fit to the data.

As a second result we analyse the angles between the directions of the 12 lines. As they are clustered into three groups corresponding to the main directions of the building, we should find values close to  $0^\circ$  within a group and values close to  $90^\circ$  between lines of different groups. The results are collected in the following table. The angles between lines in the same group scatter between  $0^\circ$  and  $14.5^\circ$ , the angles between lines of different orientation differ from  $90^\circ$  between  $0^\circ$  and  $15^\circ$ . The standard deviations of the angles scatter between  $0.4^\circ$  and  $8.3^\circ$ , this is why none of the deviations from  $0$  or  $90^\circ$  are significant.

The statistical analysis obviously makes the visual impression objective.

$l_{\min} \setminus \#$	1	2	3	4	5	6	7	8	9	10	11	12
173 [pel]	—	2.6°	2.7°	3.0°	88.6°	89.0°	88.7°	76.5°	86.7°	87.0°	86.6°	85.2°
155 [pel]	0.7	—	0.7°	1.6°	89.9°	89.7°	87.3°	75.1°	89.0°	89.2°	88.9°	87.4°
72 [pel]	0.7	0.	—	0.9°	89.4°	89.8°	87.9°	75.6°	89.4°	89.6°	89.3°	87.8°
62 [pel]	1.2	0.3	0.1	—	88.5°	88.9°	88.7°	76.4°	89.8°	90.0°	89.7°	88.2°
232 [pel]	0.6	0.0	0.1	0.2	—	0.4°	2.8°	15.3°	89.6°	89.3°	89.7°	89.1°
153 [pel]	0.3	0.1	0.0	0.1	0.1	—	2.4°	14.9°	89.5°	89.2°	89.6°	89.1°
91 [pel]	0.5	0.8	0.4	0.2	0.8	0.6	—	12.5°	89.0°	88.7°	89.1°	88.6°
113 [pel]	1.0	1.1	1.1	0.9	1.1	1.1	0.8	—	87.0°	86.7°	87.2°	87.0°
190 [pel]	1.6	0.4	0.3	0.1	0.2	0.3	1.0	1.3	—	0.4°	0.2°	1.6°
82 [pel]	1.4	0.3	0.2	0.0	0.4	0.4	1.2	1.4	0.5	—	0.5°	1.8°
103 [pel]	1.6	0.5	0.4	0.2	0.2	0.2	0.9	1.2	0.3	0.6	—	1.6°
225 [pel]	2.4	1.1	1.1	1.0	0.5	0.5	1.3	1.5	3.2	3.6	4.0	—

**Fig. 4.** Result of determining 12 lines from Fig. 3. Left column: minimal length  $l$  of the three line segments involved, upper right triangle: angels between the lines. Lower left triangle: test statistic for the deviation from 0° or 90°.

### 3 Conclusions and Outlook

We developed a rigorous estimation scheme for all types of entities in projective spaces, especially points, lines and planes in 2D and 3D, together with the corresponding transformations. The estimation requires only the minimum number of parameters for each entity, thus avoids the redundancy of the homogeneous representations. Therefore no additional constraints are required to enforce the normalization of the entities, or to enforce additional constraints such as the Plücker constraints for 3D lines. In addition we not only obtain a minimal representation for the uncertainty of the geometric elements, but also simple means to determine the Mahalanobis distance between two elements, which may be used for testing or for grouping. In both cases the estimation requires the covariance matrices of the observed entities to be invertible, which is made possible by the proposed reduced representation. We demonstrated the superiority of rigorous setup compared to a suboptimal classical method of determining vanishing points, and proved the rigour of the method with the estimation of 3D lines from points or planes. The convergence properties when using the proposed reduced representation does not change as the solutions steps are algebraically equivalent. The main advantage of the proposed concept is the ability to handle elements at or close to infinity without loosing numerical stability and that it is not necessary to introduce additional constraints to enforce the geometric entities to lie on their manifold. The concept certainly can be extended to higher level algebras, such as the geometric or the conformal algebra (see [10]) where the motivation to use minimal representations is even higher than in our context.

Acknowledgement. I acknowledge the valuable recommendations of the reviewers.

### References

1. Åstrom, K.: Using Combinations of Points, Lines and Conics to Estimate Structure and Motion. In: Proc. of Scandinavian Conference on Image Analysis. (1998)

2. Bartoli, A., Sturm, P.: Nonlinear estimation of the fundamental matrix with minimal parameters. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **26** (2004) 426–432
3. Bartoli, A., Sturm, P.: Structure-from-motion using lines: Representation, triangulation and bundle adjustment. *Computer Vision and Image Understanding* **100** (2005)
4. Begelfor, E., Werman, M.: How to put probabilities on homographies. *IEEE Trans. Pattern Anal. Mach. Intell.* **27** (2005) 1666–1670
5. Bohling, G. C., Davis, J. C., Olea, R. A., Harff, J.: Singularity and Nonnormality in the Classification of Compositional Data. *Int. Assoc. for Math. Geology*, **30**, 1, 1998, 5–20
6. Bregler, C., Malik, J.: Tracking people with twists and exponential maps. *Proc. Computer Vision and Pattern Recognition*(1998), 8–15
7. Collins, R.: Model Acquisition Using Stochastic Projective Geometry. PhD thesis, Department of Computer Science, University of Massachusetts (1993) Also published as UMass Computer Science Technical Report TR95-70.
8. Criminisi, A.: Accurate Visual Metrology from Single and Multiple Uncalibrated Images. *Distinguished Dissertations*. Springer, London, Berlin, Heidelberg (2001)
9. Förstner, W.: Optimal Vanishing Point Detection and Rotation Estimation of Single Images of a Legolandscene. *Int. Archives for Photogrammetry and Remote Sensing*, Vol. XXXVIII/3A. Paris 2010
10. Gebken, C.: Conformal Geometric Algebra in Stochastic Optimization. PhD Thesis, University of Kiel, Institute of Computer Science (2009)
11. Heuel, S.: Uncertain Projective Geometry: Statistical Reasoning For Polyhedral Object Reconstruction (Lecture Notes in Computer Science). Springer-Verlag New York, Inc., Secaucus, NJ, USA (2004)
12. Jain, A., Kurz, C., Thormälen, T., Seidel, H. P.: Exploiting Global Connectivity Constraints for Reconstruction of 3D Line Segments from Images. In: *Conference on Computer Vision and Pattern Recognition*. (2010)
13. Kanatani, K.: Statistical Optimization for Geometric Computation: Theory and Practice. Elsevier Science (1996)
14. Koch, K.R.: Parameter Estimation and Hypothesis Testing in Linear Models. Springer (1999)
15. Liebowitz, D.: Camera Calibration and Reconstruction of Geometry from Images. PhD thesis, University of Oxford, Dept. Engineering Science (2001).
16. McGlone, C.J., Mikhail, E.M., Bethel, J.S.: Manual of Photogrammetry. Am. Soc. of Photogrammetry and Remote Sensing (2004)
17. Meidow, J., Beder, C., Förstner, W.: Reasoning with uncertain points, straight lines, and straight line segments in 2D. *International Journal of Photogrammetry and Remote Sensing* **64** (2009) 125–139
18. Pottmann, H., Wallner, J.: Computational Line Geometry. Springer, Heidelberg (2010)
19. Rosenhahn, B., Granert, O., Sommer, G.: Monocular pose estimation of kinematic chains. In Dorst, L., Doran, C., Lasenby, J., eds.: *Applications of Geometric Algebra in Computer Science and Engineering*, Proc. AGACSE 2001, Cambridge, UK, Birkhäuser Boston (2002) 373–375
20. Sturm, P., Gargallo, P.: Conic fitting using the geometric distance. In: *Proceedings of the Asian Conference on Computer Vision*, Tokyo, Japan. Volume 2., Springer (2007) 784–795