On Weighting and Choosing Constraints for Optimally Reconstructing the Geometry of Image Triplets

Wolfgang Förstner

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

Abstract. Optimally reconstructing the geometry of image triplets from point correspondences requires a proper weighting or selection of the used constraints between observed coordinates and unknown parameters. By analysing the ML-estimation process the paper solves a set of yet unsolved problems: (1) The minimal set of four linearly independent trilinearities (Shashua 1995, Hartley 1995) actually imposes only three constraints onto the geometry of the image triplet. The seeming contradiction between the number of used constraints, three vs. four, can be explained naturally using the normal equations. (2) Direct application of such an estimation suggests a pseudoinverse of a 4×4-matix having rank 3 which contains the covariance matrix of the homologeous image points to be the optimal weight matrix. (3) Instead of using this singluar weight matrix one could select three linearily dependent constraints. This is discussed for the two classical cases of forward and lateral motion, and clarifies the algebraic analyis of dependencies between trilinear constraints by Faugeras 1995.
Results of an image sequence with 800 images and an Euclidean parametrization of the trifocal tensor demonstrate the feasibility of the approach.

1 Motivation and Problem

Image triplets reveal quite some advantage over image pairs for geometric image analysis. Though the geometry of the image triplet is studied quite well, implementing an *optimal* estimation procedure for recovering the orientation and calibration of the three images from point, and possibly line, correspondencies still has to cope with a number of problems.

1.1 The Task

This paper discusses the role of the trilinear constraints between observed coordinates and unknown parameters $[12, 13, 2, 8, 16]$ within an optimal estimation process for the orientation of the image triplet and shows an application within image sequence analysis.

The task formally can be described as following. We assume to have observed J sets $(P'(x', y'), P''(x'', y''), P'''(x''', y'''))_j, j = 1, ..., J$ of corresponding points

D. Vernon (Ed.): ECCV 2000, LNCS 1843, pp. 669−684, 2000.

in an image triplet. For each set of six coordinates $\boldsymbol{y}_j = (x', y', x'', y'', x''', y''')_j^\mathsf{T}$ of three corresponding points we have a set of G_j generally nonlinear constraints, here the trilinear constraints $\boldsymbol{g}_j(\boldsymbol{y}_j, \boldsymbol{\beta}) = \mathbf{0}$, which link the observed coordinates with the U parameters $\boldsymbol{\beta}$ of the orientation of the image triplet, specifically the $U = 27$ elements [16] of a $3 \times 3 \times 3$ tensor, termed trifocal tensor by [8]. There may be additional H constraints $\boldsymbol{h}(\boldsymbol{\beta}) = \mathbf{0}$ on the parameters alone, which in our case reduce the number of degrees of freedom of the trifocal tensor to 18 [8, 21, 3]. The task is to find optimal estimates for the parameters taking the uncertainty of the observed coordinates, e. g. captured in a covariance matrix $\boldsymbol{\Sigma}_{yy}$, into account.

In this work, we are primarily interested in the optimal determination of the orientation and calibration of the three cameras, not in the elements of the trifocal tensor per se. We also assume some approximate values for the parameters to be known either by the camera setup, as e. g. in motion analysis or by some direct solutions. This is no severe restriction, as such techniques are available for a large class of setups. However, the optimal estimation of the orientation and calibration parameters, though used in [21], has not been treated in depth up to now.

1.2 Problems

There is a set of yet unsolved problems which are sketched here but worked out later:

P1: The number G_j of constraints: Shashua [16] showed that there exists a set of $G_j = 9$ constraints $\boldsymbol{g}_j$ with unique properties: They are linear in the coordinates of the three homologeous points and in the elements of the trifocal tensor. Up to four of them are linearily independent. However, as six coordiates are used to determine the three coordinates of the 3D-point only three of them actually constrain the orientation of the image triplet. Therefore the number of constraints to be used should be $G_j = 3$. Thus there seems to be a *contradiction* in counting independent constraints.

P2: Choosing G_j constraints: As the choice of these, three or four constraints depends on the numbering of the images we alltogether have 12 constraints. In addition we also could use the 3 epipolar constraints, being bilinear in the coordinates, for constraining the orientation. Though the algebraic relations between these constraints are analysed in [2], no generally valid rule is known how to select constraints. Therefore we have the problem to *choose* a small subset of $G_j = 3$ constraints from a total of 15, for determination of the orientation, thus, presuming problem P1 has been clarified. The problem is non trivial because a subset, which is well suited in one geometric situation may be unfavorable in another, leading to *singularities*.

P3: Weighting the constraints: Another way to look at the problem is to ask for the optimal weighting of the constraints, being more general than choosing [15]. Then the question arizes where to obtain the weights from, how to take the geometry into account, how to deal with singular cases and how to integrate the uncertainty of the matching procedure.

P4: Modelling and optimally estimating the geometry: There are several possibilities to *model* the geometry of the image triplet: (a) using the unconstrained $U = 27$ elements of the trifocal tensor as unknown parameters, (b) using the $U = 27$ elements of the trilinear tensor as parameters with $H = 9$ constraints on the parameters, (c) using a minimal parametrization with $H = 18$ parameters or (d) even restricting the geomtry to that of calibrated cameras, leading to an *Euclidean version* [13, 17] of the trifocal tensor involving $U = 11$ parameters for the relative orientation of the first two images and the 6 parameters of the orientation of the third image. The question then arises how an *optimal estimation* could be performed in each case, and how and under which conditions the estimates differ. Moreover, how are the above mentioned problems effected by the choice of the model?

We want to discuss these problems in detail.

1.3 Outline of the Paper

We first (section 2.1) present a generic model for representing parameter estimation problems. The resulting normal equation matrix, which represents the weights of the resulting parameters, can be used to analyse the quality of the result. The trilinear constraints on the observed coordinates can be interpreted geometrically (sect. 2.2) and allow a transparent visualization of the constriants within an image triplet (sect. 2.3). Based on different models for the image triplet (sect. 3.1) we discuss the number and the weighting of the contraints (sect. 3.2) and the optimal choice of the constraints for the classical cases of lateral and forward motion, leading to general selection rules (sect. 3.3). Sect. 4 presents an example on real data to prove the concept using a metric version of the trifocal tensor.

Notation: Normal vectors $\boldsymbol{x}$ and $\boldsymbol{X}$ and matrices R are given in italics, homogeneous vectors $\mathbf{x}$ and $\mathbf{X}$ and matrices P in upright letters. If necessary for clarity, stochastical variables are underscored, e. g. $\underline{x}$ being the model variable for the observed value x. True values are indicated with a tilde, e. g. $\tilde{x}$.

2 Basics

2.1 Modelling and Estimation

In this section we describe a broad class of estimation problems (cf. [20]) whose solution is obtained by solving an optimization problem of the same general form. In all cases the task is to infer the values of U non observable quantities $\tilde{\beta}_u$ from N given observations y_n fulfilling the constraints given by the geometrical, physical or other known relations. We treat these quantities as stochastic variables in order to be able to describe their uncertainty. As this takes place in our model of the actual setup, we distinguish stochastic variables $\underline{x}$ and their realizations (observed instances) x.

Modeling the Observation Process We assume that there are two vectors of unknown quantities, the N vector $\tilde{\boldsymbol{y}} = (\tilde{y}_1, ..., \tilde{y}_n, ..., \tilde{y}_N)^\mathsf{T}$, and the U vector $\tilde{\boldsymbol{\beta}} = (\tilde{\beta}_1, ..., \tilde{\beta}_u, ..., \tilde{\beta}_U)^\mathsf{T}$ participating in the G relations

$$\boldsymbol{g}(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{\beta}}) = \mathbf{0} \tag{1}$$

whose structural form is known. The values $\tilde{y}_n$ represent the true values for the observations, which according to the model are *intended* to be made. The parameters $\tilde{\beta}_u$ are assumed not to be directly observable. In our application these constraints are the trilinearities between observed image coordiantes and parameters of the geometry of the image triplet, worked out later.

In addition, it may be that the unknown parameters $\boldsymbol{\beta}$ have to fulfill certain constraints, e. g. $\boldsymbol{\beta}^\mathsf{T}\boldsymbol{\beta} = 1$. We represent these H constraints by

$$\boldsymbol{h}(\tilde{\boldsymbol{\beta}}) = \mathbf{0} \tag{2}$$

In our application these may be 9 constraints on the 27 elements of the trifocal tensor (cf. [8]), to completely model the image geometry.

We now observe randomly perturbed values $\underline{\boldsymbol{y}}$ of the unknown vector $\tilde{\boldsymbol{y}}$. We model the random perturbation as an additive random perturbation assuming the random noise vector $\underline{\boldsymbol{e}}$ is assumed to be normally distributed with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}_{ee} = \sigma^2 \boldsymbol{Q}_{ee}$

$$\underline{\boldsymbol{y}} = \tilde{\boldsymbol{y}} + \underline{\boldsymbol{e}} \qquad \underline{\boldsymbol{e}} \sim N(\mathbf{0}, \boldsymbol{\Sigma}_{ee}) = N(\mathbf{0}, \sigma^2 \boldsymbol{Q}_{ee}) \tag{3}$$

The covariance matrix $\boldsymbol{\Sigma}_{ee}$ is separated in two factors: a positive definite symmetric matrix $\boldsymbol{Q}_{ee}$, also called the cofactor matrix (cf. Mikhail & Ackermann 1976) being an *initial covariance matrix*, giving the structure of $\boldsymbol{\Sigma}_{ee}$, and the multiplicative *variance factor* σ^2 to be estimated.

This separations has two reasons: One often only knows the ratios between the variances of the different observations and under certain conditions the estimation process is independent on the variance factor. The initial covariance matrix $\boldsymbol{Q}_{ee}$ is fixed and assumed to be known. It may result from previous experiments involving the same kinds of observations involved in the current observation. The initial covariance matrix $\boldsymbol{Q}_{ee}$ contains within it the scaling of the variables, their units, and the correlation structure of the observed variables. The *variance factor* σ^2 is an unknown variable for the multiplier on the known initial covariance matrix. It will be estimated using current data.

The complete model, represented by (1), (2) and (3), is called the GAUSS-HELMERT-*model* (cf. [9])

There are various special cases of this model. The most important one is the socalled GAUSS-MARKOFF-*model*, $\tilde{\boldsymbol{y}} = \overline{\boldsymbol{g}}(\tilde{\boldsymbol{\beta}})$ (cf. [6], p. 213, [11], p. 218) where the observation process is made explicit, like in classical regression problems.

We will apply the complete model here for using the trilinear constraints on the coefficients of the trifocal tensor for estimating the relative orientation of the image triplet and especially for *analysing the ranks of the matrices involved for discussing the number of necessary constraints.*

Estimating Parameters The estimation problem we wish to solve now is: Given $\boldsymbol{y}$, estimate $\widehat{\boldsymbol{y}}, \widehat{\boldsymbol{\beta}}$, and $\widehat{\sigma}^2$ the most probable values for $\tilde{\boldsymbol{y}}$, $\tilde{\boldsymbol{\beta}}$ and $\tilde{\sigma}^2$.

We solve this problem by finding the value $(\widehat{\boldsymbol{y}}, \widehat{\boldsymbol{\beta}})$ for $(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{\beta}})$ that minimizes the weighted sum of squares of residuals, the weight matrix being the inverse covariance matrix $\phi(\boldsymbol{y}, \tilde{\boldsymbol{y}}) = 1/2\ (\boldsymbol{y} - \tilde{\boldsymbol{y}})^{\mathsf{T}} \mathsf{Q}_{yy}^{-1} (\boldsymbol{y} - \tilde{\boldsymbol{y}})$ subject to the constraints $\boldsymbol{g}(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{\beta}}) = \mathbf{0}$ and $\boldsymbol{h}(\tilde{\boldsymbol{\beta}}) = \mathbf{0}$. This is equivalent to finding the minimum of

$$\Phi(\tilde{\boldsymbol{y}}, \boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \frac{1}{2}(\boldsymbol{y} - \tilde{\boldsymbol{y}})^{\mathsf{T}} \mathsf{Q}_{yy}^{-1} (\boldsymbol{y} - \tilde{\boldsymbol{y}}) + \boldsymbol{\lambda}^T \boldsymbol{g}(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{\beta}}) + \boldsymbol{\mu}^T \boldsymbol{h}(\tilde{\boldsymbol{\beta}}) \tag{4}$$

where $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$ are G and H-vectors of Lagrangian multipliers. The solution is the ML-estimate, in case observations actually follow a normal distribution. Otherwise they are (locally) best linear unbiased estimates, i. e. estimates with smallest variance. The general solution of this optimization problem is given in the appendix.

We only need the normal equation matrix N here, which contains the *covariance matrix* $\boldsymbol{\Sigma}_{\widehat{\beta}\widehat{\beta}}$ *of the estimated unknown parameters* $\widehat{\boldsymbol{\beta}}$ in its inverse. With the Jacobians A and B of $\boldsymbol{g}$ with respect to the unknowns and the observations, and the Jacobian H of $\boldsymbol{h}$ with respect to the unknown parameters and the assumptions that these matrices have full rank we obtain the normal equation matrix

$$\mathsf{N} = \begin{pmatrix} \mathsf{A}^{\mathsf{T}}(\mathsf{B}\mathsf{Q}_{yy}\mathsf{B}^{\mathsf{T}})^{-1}\mathsf{A} & \mathsf{H}^{\mathsf{T}} \\ \mathsf{H} & \mathbf{0} \end{pmatrix} = \begin{pmatrix} \overline{\mathsf{N}} & \mathsf{H}^{\mathsf{T}} \\ \mathsf{H} & \mathbf{0} \end{pmatrix} = \sigma^{-2} \begin{pmatrix} \boldsymbol{\Sigma}_{\widehat{\beta}\widehat{\beta}} & \mathsf{S}^{\mathsf{T}} \\ \mathsf{S} & \mathsf{T} \end{pmatrix}^{-1} \tag{5}$$

with some matrices S and T, cf. (28) appendix.

We will be able to identify the rows of the matrix A with the Jacobian of the trilinear constraints w. r. t. the elements of the trifocal tensor, the matrix $(\mathsf{B}\mathsf{Q}_{yy}\mathsf{B}^{\mathsf{T}})^{-1}$ with the sought weight matrix for the trilinearities containing the (initial) covariance matrix Q_{yy} of the observed coordintes and analyse the rank of these matrices.

2.2 Projection Matrices and their Interpretation

The geometric setup of three images is given by

$$\mathbf{x}_i = \mathsf{P}_i \mathbf{X} \quad i = 1, 2, 3 \tag{6}$$

which relate the coordinates $\mathbf{X}^{\mathsf{T}} = (X, Y, Z, 1)$ of the object point to the three sets of coordiantes $\mathbf{x}_i^{\mathsf{T}} = (u_i, v_i, w_i)$ with the (Euclidean) image coordinates $x' = u_1/w_1$, $y' = v_1/w_1$, $x'' = u_2/w_2$, etc. The three projection matrices are

$$\mathsf{P}_1 = \begin{pmatrix} \mathbf{1}^{\mathsf{T}} \\ \mathbf{2}^{\mathsf{T}} \\ \mathbf{3}^{\mathsf{T}} \end{pmatrix}, \quad \mathsf{P}_2 = \begin{pmatrix} \mathbf{4}^{\mathsf{T}} \\ \mathbf{5}^{\mathsf{T}} \\ \mathbf{6}^{\mathsf{T}} \end{pmatrix}, \quad \mathsf{P}_3 = \begin{pmatrix} \mathbf{7}^{\mathsf{T}} \\ \mathbf{8}^{\mathsf{T}} \\ \mathbf{9}^{\mathsf{T}} \end{pmatrix} \tag{7}$$

where the rows are indicated with bold face numbers. With the standard parametrization of the projection matrices

$$\mathsf{P}_i = \mathsf{K}_i \mathsf{R}_i (\mathsf{I} | -\boldsymbol{X}_{oi}) \tag{8}$$

Eq. (6) relates Euclidean object-space to Euclidean image space, capturing the (Euclidean) object coordiantes $\boldsymbol{X}_o$ of the projection centre, the rotation R and the calibration K, being an upper triangular matrix with 5 free parameters.

We now describe the geometry of the image triplet using the vectors **1**, **2** etc. in detail. We use the following interpretation of the rows **1**, **2** etc. of the projection matrices (cf. [2]): In case $u_1 = 0$ and (v_1, w_1) arbitrary, we have $\mathbf{1}\cdot\mathbf{X} = 0$, thus the vector **1** represents the homogeneous coordinates of the plane passing through the y'- and the z'-axis in the first camera; in the special case of $\mathsf{K} = Diag(c_1, c_2, 1)$, i. e. reduced image coordinates but arbitrary focal lengths c_i, they are perpendicular to the x'-axis. By analogy, **4** and **7** are planes containing the $y^{(i)}$- and the $z^{(i)}$-axes in the second and the third camera, **2**, **5** and **8** are planes containing the $x^{(i)}$- and the $z^{(i)}$-axes, and **3**, **6** and **9** are planes containing the $x^{(i)}$- and the $y^{(i)}$-axes in the three cameras. Observe, all these planes pass through the corresponding projection centre.

As $u_1 : v_1 : w_1 = (\mathbf{1}\cdot\mathbf{X}) : (\mathbf{2}\cdot\mathbf{X}) : (\mathbf{3}\cdot\mathbf{X})$ and correspondingly for the other cameras, we have the following equivalent homogeneous constraints for the image coordinates:

$$\begin{pmatrix} \mathbf{A}_1^\mathsf{T} \\ \mathbf{B}_1^\mathsf{T} \\ \mathbf{D}_1^\mathsf{T} \\ \mathbf{A}_2^\mathsf{T} \\ \mathbf{B}_2^\mathsf{T} \\ \mathbf{D}_2^\mathsf{T} \\ \mathbf{A}_3^\mathsf{T} \\ \mathbf{B}_3^\mathsf{T} \\ \mathbf{D}_3^\mathsf{T} \end{pmatrix} \mathbf{X} \doteq \begin{pmatrix} u_1\mathbf{3}^\mathsf{T} - w_1\mathbf{1}^\mathsf{T} \\ v_1\mathbf{3}^\mathsf{T} - w_1\mathbf{2}^\mathsf{T} \\ u_1\mathbf{2}^\mathsf{T} - v_1\mathbf{1}^\mathsf{T} \\ u_2\mathbf{6}^\mathsf{T} - w_2\mathbf{4}^\mathsf{T} \\ v_2\mathbf{6}^\mathsf{T} - w_2\mathbf{5}^\mathsf{T} \\ u_2\mathbf{5}^\mathsf{T} - v_2\mathbf{4}^\mathsf{T} \\ u_3\mathbf{9}^\mathsf{T} - w_3\mathbf{7}^\mathsf{T} \\ v_3\mathbf{9}^\mathsf{T} - w_3\mathbf{8}^\mathsf{T} \\ u_3\mathbf{8}^\mathsf{T} - v_3\mathbf{7}^\mathsf{T} \end{pmatrix} \mathbf{X} \cong \begin{pmatrix} x'\mathbf{3}^\mathsf{T} - \mathbf{1}^\mathsf{T} \\ y'\mathbf{3}^\mathsf{T} - \mathbf{2}^\mathsf{T} \\ x'\mathbf{2}^\mathsf{T} - y'\mathbf{1}^\mathsf{T} \\ x''\mathbf{6}^\mathsf{T} - \mathbf{4}^\mathsf{T} \\ y''\mathbf{6}^\mathsf{T} - \mathbf{5}^\mathsf{T} \\ x''\mathbf{5}^\mathsf{T} - y''\mathbf{4}^\mathsf{T} \\ x'''\mathbf{9}^\mathsf{T} - \mathbf{7}^\mathsf{T} \\ y'''\mathbf{9}^\mathsf{T} - \mathbf{8}^\mathsf{T} \\ x'''\mathbf{8}^\mathsf{T} - y'''\mathbf{7}^\mathsf{T} \end{pmatrix} \mathbf{X} = \mathbf{0} \qquad (9)$$

The vectors $\mathbf{A}_i$, $\mathbf{B}_i$, $\mathbf{D}_i$ have a specific geometric meaning [18]:

The vectors $\mathbf{A}_i$ represent planes through the origin of the i-th camera, as they are linear combinations of the plane vectors; they pass through the v_i-axis of the i-th camera, as it is contained in both planes **1** and **3**; they pass through the image point P_i, due to eq. (9); therefore they intersect the image plane in the line $u_i = const$. The vectors $\mathbf{B}_i$ represent planes through the origin of the i-th camera, pass through the $x^{(i)}$-axis of the i-th camera, pass through the image point P_i and thus intersect the image plane in the line $v_i = const$. Now, the vectors $\mathbf{D_i}$ represent planes through the origin of the i-th camera, pass through the $z^{(i)}$-axis of the i-th camera, pass through the image point P_i and thus intersect the image plane radially, fixing the *direction*, motivating the notation.

Observe, the planes $\mathbf{D}_i$ are not defined or are instable for points identical or close to the origin $(0, 0)$. Thus, planes $\mathbf{A}_i$, $\mathbf{B}_i$ and $\mathbf{D}_i$ fix the $x^{(i)}$-, the $y^{(i)}$- and the 'directional' coordinate. Only two of the three constraints for each camera are independent.

2.3 Constraints between Points of an Image Triplet

Shashua's Four Constraints on the Trifocal Tensor Elements We now easily can write down Shashua's constraints [16]. They can be formulated using the above mentioned planes, by establishing quadrupels of planes which should intersect in a 3D point, which is equivalent to requiring the 4×4 matrix of the 4 plane coordinate vectors to be singular or its determinant to vanish:

$$D_1^S \doteq |\mathbf{A}_1\,\mathbf{B}_1\,\mathbf{A}_2\,\mathbf{A}_3| = 0, \quad D_2^S \doteq |\mathbf{A}_1\,\mathbf{B}_1\,\mathbf{A}_2\,\mathbf{B}_3| = 0 \tag{10}$$

$$D_3^S \doteq |\mathbf{A}_1\,\mathbf{B}_1\,\mathbf{B}_2\,\mathbf{A}_3| = 0 \quad D_4^S \doteq |\mathbf{A}_1\,\mathbf{B}_1\,\mathbf{B}_2\,\mathbf{B}_3| = 0 \tag{11}$$

where $|\cdots|$ denotes the determinant of the vectors. The point-line-line constraint results from the fact that the first two vectors $\mathbf{A}_1$ and $\mathbf{B}_1$ fix the ray through the point in the first image and the two other vectors represent lines through the points in the second and the third image each (cf. the geometric interpretation above)[1].

Observe that these constraints are linear in all image coordinates, as each of these coordinates appears only once in the determinants and the w_i-coordinate can be set to 1 for all image image points, cf. (9c).

Shashua moreover showed that the constraints (10), (11) can be written as linear functions of the 27 entries of a $3 \times 3 \times 3$ tensor with elements $\mathbf{t}$, thus each is of the form

$$D_{lj}^S = \boldsymbol{\alpha}_{lj}^{S\mathsf{T}}\,\mathbf{t} = 0 \qquad l = 1, 2, 3, 4 \tag{12}$$

where the 27-vector $\boldsymbol{\alpha}_{lj}^S \doteq \boldsymbol{\alpha}_l^S(\boldsymbol{y}_j)$ only depends on the six coordinates of the point triple collected in the 6-vector $\boldsymbol{y}_j$ (here indexed with j to indicate the used point triple), and the 27-vector $\mathbf{t}$ contains the tensor coefficients. Shashua showed the 27×4 matrix

$$\mathsf{A}_j^S = (\boldsymbol{\alpha}_{1j}^S, \boldsymbol{\alpha}_{2j}^S, \boldsymbol{\alpha}_{3j}^S, \boldsymbol{\alpha}_{4j}^S) = \left(\frac{\partial D_{lj}^S}{\partial t_k}\right)^{\mathsf{T}} \quad \begin{matrix} k = 1, ..., 27;\ j = 1, ..., J \\ l = 1, ...4 \end{matrix} \tag{13}$$

to have rank four. Observe that A_j^S is the transposed Jacobian of the constraints (10 ff.) with respect to the parameters $\mathbf{t}$.

This suggests 4 constraints are necessary if one wants to exploit the full information of the image points for recovering the geometry of the image triplet.

Three Constraints between the Observations and the Triplet's Geometry However, we could argue only to need three constraints:

If one solves the basic projection equations (6) for the 6 observed coordinates one obtains 6 inhomogeneous equations. One now can take three of them and solve for the 3 coordinates of the object point. Substituting these object coordinates into the other three inhomogeneous equations yields *three* constraints between the six imge coordinates and the parameters of the geometry of the

[1] The four constraints correspond to those given in [8]: $\sum_k u'_k(u''_i u'''_j T_{k33} - u'''_j T_{ki3} - u''_i T_{k3j} + T_{kij}) = 0$ with the combinations $(1,1), (1,2), (2,1), (2,2)$, for the indices i and j and homogeneous coordinates (u'_1, u'_2, u'_3) and $u'_3 = 1$ etc.

image triplet, *independent* on the parametrization. Thus there can only be 3 independent constraints between the observed image coordiantes and the parameters of the geometry of the image triplet.

In a general setup one could argue that 1.) between the first two points P_1 and P_2 the epipolar constraints should be valid and that 2.) the 3D point, determined from the first two images, should map into the third image. This prediction was the basis for the derivation of the trifocal tensor in [8].

The epipolar constraint then reads as[2]

$$D_1^I \doteq |\mathbf{A}_1\,\mathbf{B}_1\,\mathbf{A}_2\,\mathbf{B}_2| = 0 \tag{14}$$

The first two vectors $\mathbf{A}_1$ and $\mathbf{B}_1$ span the ray in the first image, whereas the last two vectors span the the ray in the second image, which should intersect. The 3D point from the first two images could be determined as the intersection of the planes $\mathbf{A}_1$, $\mathbf{B}_1$ and $\mathbf{A}_2$ which should ly in the two planes $\mathbf{A}_3$ and $\mathbf{B}_3$, which gives rise to two further constraints, namely:

$$D_2^I \doteq |\mathbf{A}_1\,\mathbf{B}_1\,\mathbf{A}_2\,\mathbf{A}_3| = 0, \quad D_3^I \doteq |\mathbf{A}_1\,\mathbf{B}_1\,\mathbf{A}_2\,\mathbf{B}_3| = 0 \tag{15}$$

which are identical to the first two D_1^S and D_2^S of Shashua's constraints[3].

Singular Cases Unfortunately this set of constraints does not work in general.

First, assume the three images have collinear projection centres, establishing the X-axis in 3D and the rotation matrices are $R_i = I$. Then the two planes $\mathbf{A}_1$ and $\mathbf{A}_2$ intersect in a line parallel to the Y-axis, which, when intersected with $\mathbf{B}_1$ yields a well defined 3D point.

Now, if the three projection centres establish the Y-axis the two planes $\mathbf{A}_1$ and $\mathbf{A}_2$ are identical, as they are epipolar planes. Thus the 3D point *cannot* be determined using these two planes. In case the constraints D_2^I and D_3^I would be replaced by the last two constraints (11) of Shashua, we would be able to determine and predict the 3D-point in this case, but not in the previous one.

We therefore need to clarify the number of necessary constraints and discuss the selection or, more general, the weighing of the constraints.

3 Constraints within the Estimation Process

We now want to establish a statistical interpretation of such dependencies. Therefore we follow sect. (2.1), and model the reconstruction of the geometry of the image triplet.

3.1 Models

We distinguish three parametrizations:

M1: Tensor coefficients: We use the 27 elements $\mathbf{t}$ of the trifocal tensor as parameters to desrcibe the geometry. We therefore use (12) as constraints for each point triplet. We have to distinguish this model from the following:

[2] The superscript I indicates case I in the analysis later.

[3] and to Hartley's constraints with indices $(1,1)$ and $(1,2)$ cf. previous footnote .

M2: Projective parametrization: We use a minimal parametrization of the trifocal tensor with 18 parameters (cf. e. g. [21]). This leads to a projective reconstruction. We need not specify the parametrization for our analysis. Instead, we also could use the U=27 tensor coefficients and H=9 appropriate constraints between these parameters (cf. [3]).

M3: metric parametrization: We use a metric parametrization of the trifocal tensor with only 11 parameters in order to achieve an Euclidean reconstruction. The reason is: in our special application of image sequence analysis, we are able to calibrate the cameras in beforehand. Therefore we only have 11 parameters to specify the geometry of the image triplet, namely the 5 parameters of the relative orientation of the first two cameras as above and the 6 parameters of the exterior orientation of the third camera (cf. [12, 13, 17]). In our implementation we actually parametrize the orientation by the two translation vectors $\boldsymbol{X}_{o2}$ and $\boldsymbol{X}_{o3}$, and the two quaterions $\mathbf{q}_2$ and $\mathbf{q}_3$ for the rotations, fixing $\boldsymbol{X}_{o1} = \mathbf{0}$, and yielding $U = 14$ parameters with the $H = 3$ constraints $\tilde{\boldsymbol{X}}_{o2}^{\mathsf{T}} \tilde{\boldsymbol{X}}_{o2} = 1$, $\tilde{\mathbf{q}}_2^{\mathsf{T}} \tilde{\mathbf{q}}_2 = 1$ and $\tilde{\mathbf{q}}_3^{\mathsf{T}} \tilde{\mathbf{q}}_3 = 1$. This model will be used in the example.

In the last two cases M2 and M3 we may use the same constraints as above, by just replacing the 27 elements t_k of the trifocal tensor by 27 functions $t_k(\boldsymbol{\beta})$ of the 18 and 14 unknown parameters, thus the constraints (12) now read as

$$g_{lj}^S(\tilde{\boldsymbol{y}}_j, \tilde{\boldsymbol{\beta}}) = D_l^S(\tilde{\boldsymbol{y}}_j, \tilde{\boldsymbol{\beta}}) = \boldsymbol{\alpha}_l^S(\tilde{\boldsymbol{y}}_j)^{\mathsf{T}}\, \mathbf{t}(\tilde{\boldsymbol{\beta}}) = 0 \qquad l = 1, 2, 3, 4 \tag{16}$$

The corresponding contraints of set A (14), (15) read as:

$$g_{lj}^A(\tilde{\boldsymbol{y}}_j, \tilde{\boldsymbol{\beta}}) = D_l^A(\tilde{\boldsymbol{y}}_j, \tilde{\boldsymbol{\beta}}) = \boldsymbol{\alpha}_l^A(\tilde{\boldsymbol{y}}_j)^{\mathsf{T}}\, \mathbf{t}(\tilde{\boldsymbol{\beta}}) = 0 \qquad l = 1, 2, 3 \tag{17}$$

In case of model 3 we in addition have the $H = 3$ constraints between the parameters only:

$$\boldsymbol{h}^{\mathsf{T}}(\tilde{\boldsymbol{\beta}}) = (\tilde{\boldsymbol{X}}_{o2}^{\mathsf{T}} \tilde{\boldsymbol{X}}_{o2} - 1 \quad \tilde{\mathbf{q}}_2^{\mathsf{T}} \tilde{\mathbf{q}}_2 - 1 \quad \tilde{\mathbf{q}}_3^{\mathsf{T}} \tilde{\mathbf{q}}_3 - 1) = \mathbf{0} \tag{18}$$

3.2 Number and Weighting of Constraints

We now discuss the left upper submatrix $\overline{N}$ from (5) in our context. In case of $j = 1, ..., J$ statistically independent triplets of points, thus $Q_{yy} = Diag(Q_{y_j y_j})$, which is no restriction in practical cases, it can be written as

$$\overline{N} = \sum_{j=1}^{J} \overline{N}_j = \sum_{j=1}^{J} A_j W_j A_j^{\mathsf{T}} = \sum_{j=1}^{J} A_j (B_j^{\mathsf{T}} Q_{y_j y_j} B_j)^{-1} A_j^{\mathsf{T}} \tag{19}$$

using $A = (A_j^{\mathsf{T}})$ and $B = (B_j^{\mathsf{T}})$.

Each part $\overline{N}_j$ depends on three matrices, A_j, B_j and $Q_{y_j y_j}$. They have a very specific semantics. They give the key to the solution of the stated problems:
Coefficient matrix A_j: The matrix A_j is the Jacobian of the constraints $\boldsymbol{g}_j(\boldsymbol{y}_j, \boldsymbol{\beta})$ with repect to the unknown parameters evaluated at the fitted values $\widehat{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{y}}_j$ of the parameters and the observations resp. (cf. App.).

In case the constraints are linear in the unknown parameters the matrix A_j only depends on the fitted coordinates $\widehat{\boldsymbol{y}}_j$. Moreover, then one may use them for a direct solution of the unknowns $\boldsymbol{\beta}$ being the eigenvector corresponding to the smallest eigenvalue of $\dot{\mathsf{N}} = \mathsf{A}^\mathsf{T}\mathsf{A} = \sum_j \mathsf{A}_j\mathsf{A}_j^\mathsf{T}$ minimizing the algebraic distance. This shows the close relation between the optimal nonlinear estimation and the direct solution: The constraints are not weighted; the direct solution obviously is an approximation. The weights W_j depend on B_j, which itself depends on the unknown parameters, thus are not available in a one-step solution. Due to the linear independency of the 4 constraints per point triplet, at least 7 points are necessary (cf. [16]) for the determination of the 27 tensor elements.

In case of 18 parameters the Jacobian A_j turns out have rank 3 in general, as can be shown using MAPLE. This is due to the projection of the 27 dimensional space of tensor parameters $\boldsymbol{t}$ to the 18 dimensional space of parameters $\boldsymbol{\beta}$.

This can be geometrically visualized as follws: Without posing restrictions, assume the translation vector $\boldsymbol{X}_{o2} = (X_{o2}, 0, 0)^\mathsf{T}$, and calibrated cameras with $\mathsf{K}_i = \mathsf{I}$, $\mathsf{R}_i = \mathsf{I}$. Then the two last constraints D_3^S and D_4^S *both* constrain the two first rays to follow the *epipolar geometry*, if the object point is in general position: this is because, $\mathbf{A}_1$ and $\mathbf{B}_1$ and the last plane $\mathbf{A}_3$ or $\mathbf{B}_3$ in (10) fix the object point. The plane $\mathbf{B}_2$ determined by the y_2-coordinate then has to pass through that point, in all three cases yielding the same constraint $y'' - y' = 0$.

Analytically, the two constraints in general are polynomials, which factor into, say, u_3v_3 and u_4v_4, where *in general position of the point* the first factors u_3 and u_4 are non zero and the second factors are identical, $v_3 \equiv v_4$, thus *both constraints, though algebraically different, impose the same restrictions onto the image geometry.*

This shows the two geometric setups, with 27 and 18 parameters resp., to differ in essence, solving problem P1, and explains why there is no real contradiction between the number of necessary constraints: SHASHUA's set of 4 constraints is necessary for estimating the geometry coded in the elements of the trifocal tensor, whereas only 3 constraints are necessary in case one wants to determine the projective geometry of the image triplet with 18 parameters. Observe, in this case one also could take the 27 elements of the trifocal tensor as unknowns $\boldsymbol{\beta}$ and introduce 9 constraints $\boldsymbol{h}$ on these parameters alone, this would not change the reasoning.

Coefficient matrix B_j: The matrix B_j is the Jacobian of the constraints with repect to the observations evaluated at the fitted values $\widehat{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{y}}_j$ of the parameters and the observations resp. (cf. App.).

It is implicitely used in the solution to determine the (preliminary) *covariance matrix* $\mathsf{Q}_{g_jg_j} = \mathsf{B}_j^\mathsf{T}\mathsf{Q}_{y_jy_j}\mathsf{B}_j$ of the contadictions $\boldsymbol{c} = \boldsymbol{g}_j(\underline{\boldsymbol{y}}_j, \beta^{(0)}) \neq \mathbf{0}$, i. e. the deviation of the constraint evaluated at the observations $\underline{\boldsymbol{y}}_j$ and the approximate values $\boldsymbol{\beta}^{(0)}$ of the parameters by error propagation. The weightmatrix $\mathsf{W}_j = \mathsf{Q}_{g_jg_j}^{-1}$, being the inverse of this covariance matrix, therefore is the optimal choice. This solves problem P2, namely the choice of the weight matrix.

If model M1 with the 27 tensor coefficients as unknowns is chosen, the rank of this weight matrix in general is four, indicating that all 4 constraints actually are relevant and can be adequately weighted. However, for models M2 and M3 with 18 or less parameters describing the projective or Euklidean geometry the weightmatrix in general has rank 3. This confirms the fact that only three independent constraints are available. The Jacobians A_j and B_j have the same null space. Taking some generalized inverse

$$\mathsf{W}_j = \mathsf{Q}^-_{g_j g_j} = (\mathsf{B}_j^\mathsf{T} \mathsf{Q}_{y_j y_j} \mathsf{B}_j)^- \tag{20}$$

when using more than 3 constraints does not lead to a different solution of the estimation problem.

This type of weighting with W_j has been used by [21]. The Jacobian B corresponds to Jacobian J in their eq. (20), which they state to have rank 3. They use the pseudo inverse $(\mathsf{J}\boldsymbol{\Sigma}_x^{1,2,3}\mathsf{J}^\mathsf{T})^+$ (24) (via a SVD) instead of the normal inverse $(\mathsf{B}_j\boldsymbol{\Sigma}_{yy}\mathsf{B}^\mathsf{T})^{-1}$ of a minimal set. This is more time consuming, compared to inverting a regular 3×3-matrix, especially if the number of used constraints is much larger than 3, as this computation has to be performed for every point triple in every iteration, which may be essential in real time applications.

However, the analysis confirms the direct 6-point solution of [21]to be a solution for the minimal number of points, as the number of free parameters of the trifocal tensor is 18 [16].

The weighing proposed in [15] is only an approximation as the rank of the weight matrix there is 4 instead of 3.

Covariance matrix $\mathsf{Q}_{y_j y_j}$: As to expected, the weighting of the constraints depends on the uncertainty of the feature points or generally of the matching procedure. This uncertainty can be captured in the covariance matrix $\mathsf{Q}_{y_j y_j}$ of the 6 coordiantes. Usually a diagonal matrix I will be sufficient. If the matching technique provides a realistic internal estimate of the variances this could be used to improve the result.

Observe, if $\mathsf{Q}_{y_j y_j} = \mathsf{I}$ then the direct solution would use the smallest eigenvector of $\mathsf{A}^\mathsf{T}(\mathsf{B}\mathsf{B}^\mathsf{T})^{-1}\mathsf{A}$. This *least squares solution* is identical to that given by [19], as $(\mathsf{B}\mathsf{B}^\mathsf{T})^{-1} = Diag(1/|\nabla g_j|^2)$ with the gradient magnitude of the constraints w. r. t. the observations. However, it here naturally follows from the general solution in a statistical estimation framework as a special case, and shows how to handle observed quantities which are *correlated.*

3.3 Choosing Independent Constraints

Instead of using a pseudo inverse for automatically getting the correct weight we also could choose a set of three independent constraints. The chosen set obviously will depend on the position of the object point with respect to the trifocal plane: If it is off the trifocal plane, three pairs of epipolar constraints would work. Thus we only analyse the important case where the projections centres are collinear. Then *all* object points lie on *a* trifocal plane, requiring at least one trilinear constraints on the tensor coefficients. We summarize the analysis from [5] here.

We assume image sequences with $\mathsf{R}_i^{(0)} = \mathsf{I}$, $\mathsf{K}_i^{(0)} = Diag(c, c, 1)$, thus principal distance $c \doteq c^{(0)}$ and distinguish *forward motion* in Z-direction with $\boldsymbol{X}_{o2}^{(0)} = (0, 0, B)^\mathsf{T}$, $\boldsymbol{X}_{o3}^{(0)} = (0, 0, 2B)^\mathsf{T}$ and *lateral motion* in X-direction with

$\boldsymbol{X}_{o2}^{(0)} = (B, 0, 0)^\mathsf{T}$, $\boldsymbol{X}_{o3}^{(0)} = (2B, 0, 0)^\mathsf{T}$, thus base length $B = B^{(0)}$. We apply two different sets of constraints. The first is set A as in eq. (14) and (15). Trying to obtain full symmetry by using every coordinate twice and fixing each image ray in one of the three constraints [2, 18]) we obtain constraint set II:

$$D_1^{II} = |\mathbf{A}_1\, \mathbf{B}_1\, \mathbf{A}_2\, \mathbf{B}_3| = 0,\ D_2^{II} = |\mathbf{B}_1\, \mathbf{A}_2\, \mathbf{B}_2\, \mathbf{A}_3| = 0,\ D_3^{II} = |\mathbf{A}_1\, \mathbf{B}_2\, \mathbf{A}_3\, \mathbf{B}_3| = 0$$

We give the determinants of the matrices $\mathsf{Q}_{g_j g_j} = (\boldsymbol{B}^\mathsf{T} \mathsf{Q}_{y_j y_j} \mathsf{B})$, being proportional to the corresponding covariance matrices, in dependency of the object coordiantes (X, Y, Z) for lateral (l) and forward (f) motion and for set I and II, $d(Z)$ being a function of Z only:

$$|\mathsf{Q}_{g_j g_j}^{(l,I)}| = 18B^2c^6 \qquad |\mathsf{Q}_{g_j g_j}^{(f,I)}| = X^4(X^2+Y^2)\cdot d(Z) \tag{21}$$

$$|\mathsf{Q}_{g_j g_j}^{(l,II)}| = 0 \qquad |\mathsf{Q}_{g_j g_j}^{(f,II)}| = X^2Y^2(X^2+Y^2)\cdot d(Z) \tag{22}$$

Only if the determinat is not 0 the weightmatrix W_j has the proper rank. Therefore, the set I obviously is useful for all points in lateral motion, as the covariance matrix is regular, with a determinant *independent on the position*. The symmetric set II, however, is not useful at all in lateral motion. This is plausible, as only the y-coordinates are taken into account, i. e. this set then is a variation of the trifold use of the epipolar constraint. Both sets do quite a good job in forward motion, however lead to singularities if the points lie on the axes, on the x-axis for set I, on one of both for set II. Observe, that the origin $(0, 0)$ is the focus of expansion (FOE): points in the direction of the motion cannot be used at all, which is counter intuitive. They actually only constrain the rotation, not the translation, thus lead to only two constraints, causing the rank deficiency.

General rules for *choosing three constraints* are the following, solving problem P2 while distinuishing between 1, 2 and 3 trilinear constraints within the set:

1. One trilinear constraint and two epipolar constraints: The trilinear constraint in lateral motion needs to be one of $|\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \mathbf{B}_i|$ $i = 1, 2, 3$. In forward motion we distinguish between points right or left of the FOE, for which the previous constraints works, and points above or below the FOE, for which one chooses one of $|\mathbf{B}_1, \mathbf{B}_2, \mathbf{B}_3, \mathbf{A}_i|, i = 1, 2, 3$[4].
2. Two trilinear constraints and one epipolar constraint: For lateral motion (X-direction) choose set I. For forward motion we again choose the sets according to position relative to the FOE, namely the determinants
$$D_1^{l,r} = |\mathbf{A}_1\, \mathbf{B}_1\, \mathbf{A}_2\, \mathbf{B}_2|,\ D_2^{l,r} = |\mathbf{A}_1\, \mathbf{B}_1\, \mathbf{A}_2\, \mathbf{A}_3|,\ D_3^{l,r} = |\mathbf{A}_1\, \mathbf{B}_1\, \mathbf{A}_2\, \mathbf{B}_3|$$
$$D_1^{a,b} = |\mathbf{A}_1\, \mathbf{B}_1\, \mathbf{A}_2\, \mathbf{B}_2|,\ D_2^{a,b} = |\mathbf{A}_1\, \mathbf{B}_1\, \mathbf{B}_2\, \mathbf{A}_3|,\ D_3^{a,b} = |\mathbf{A}_1\, \mathbf{B}_1\, \mathbf{B}_2\, \mathbf{B}_3|$$
to be zero (*l,r* = left/right, *a,b* = above/below the FOE).
3. Three trilinear constraints *with the same ray fixed* in all constraints are generally independent *if* no constraint contains 3 planes parallel to a/the trifocal plane[5]. E. g. the set $D_1 = |\mathbf{A}_1, \mathbf{B}_1, \mathbf{A}_2, \mathbf{A}_3| = 0$, $D_2 = |\mathbf{A}_1, \mathbf{B}_1, \mathbf{A}_2, \mathbf{B}_3| = 0$, $D_3 = |\mathbf{A}_1, \mathbf{B}_1, \mathbf{B}_2, \mathbf{A}_3| = 0$ is independent in general; in lateral motion for all points, in forward motion at all points except with $X = 0$ or $Y = 0$.

[4] [13] proposes $|\mathbf{3}, \mathbf{D}_1, \mathbf{D}_2, \mathbf{D}_3| = 0$ in lateral motion only useful for points with $X \neq 0$.

[5] The set of constraints discussed in [2], p. 16 $T_{1,2,3,5} = |\mathbf{D}_1, \mathbf{B}_1,\ \mathbf{D}_2,\ \mathbf{D}_3|$, $T_{1,3,4,5} = |\mathbf{D}_1, \mathbf{D}_2,\ \mathbf{B}_2,\ \mathbf{D}_3|$, $T_{1,3,5,6} = |\mathbf{D}_1, \mathbf{D}_2,\ \mathbf{D}_3,\ \mathbf{B}_3|$ has rank 1 in lateral motion in x- or

4 Example

The usefulness of the estimation procedure for the metric version of the trifocal tensor is investigated [1]. An image sequence with 5300 images is taken from a car. The measured speed is shown in figure 1 top. The camera is looking ahead, establishing the case *forward motion*. A subsecquence of 800 images has been evaluated w. r. t. the geometric analysis of image pairs and image triplets. A subsection of 100 images was finally used for a bundle triangulation.

The Procedure: After initializing the procedure, interest points are selected in image i which promise good correspondence [4]. Using the correspondencies from the two previous frames we predict points in the current image using two trilinear constraints suited for that point. Thereby we assume *constant motion*, thus constant $\boldsymbol{X}_o$ and R. All interest points within an adaptive search area are checked for consistency using normalized crosscorrelation. Possibly their position is corrected based on the point in image $i-1$ using a least squares matching procedure [7], chap. 16, at the same time yielding internal estimates for the uncertainty $\boldsymbol{\Sigma}_{y_j y_j}$.

These point triplets are used for estimation. We applied the set of constraints I (14, 15) here and used a pseudo inverse to cope with singularities, which are possible (cf. (21b)). The 14 parameters with the 3 constraints of the third model are estimated using the GAUSS-HELMERT-model, however in a robustified version, by a reweighting scheme following [10]. Figure 1 bottom shows the number of used points per successfully determined image triplet, which excludes the images with velocity 0.

Finally, new points are detected and possibly linked to the previous image. The next image $i+1$ is taken as third in the next image triplet, which uses the *metric parametrization* of the two previous images as approximate values. This chaining is not meant to be optimal not even consistent, as it only is used to yield approximate values for the image sequence, which then were to be optimally reconstructed in one process using a bundle adjustment.

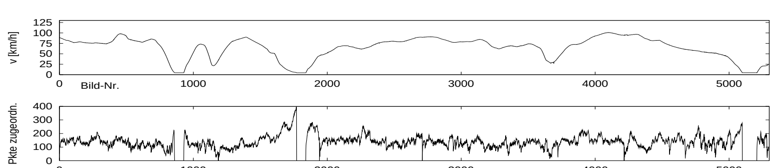

Fig. 1. *shows speed (top) and number of matched points over 5300 images (bottom)*

Results: Some results of the extensive experiments, documented in [1] can be summarized as follows:

y-direction, as the three planes $\mathbf{D}_i$ are parallel to the z-axis and should intersect in one ray, which is expressed equivalently by all three constraints; the set has rank 0 in forward motion in z-direction as they all contain three planes passing through the motion axis.

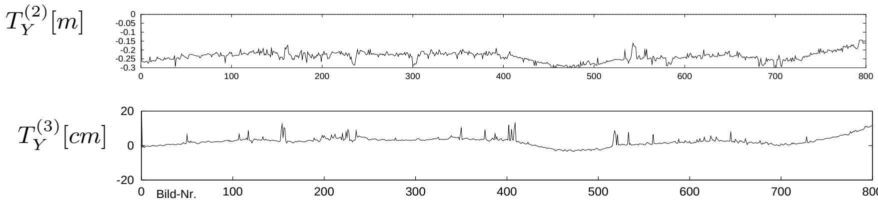

Fig. 2. *shows above the Y-coordinate of the translation vector* $\boldsymbol{T}_2$ *determined from the first 800 image pairs, revealing quite a number of erroneous values cause by mismatches. The lower row shows the Y-coordinate determined from image triplets, clearly demonstrating the effect of higher reliability from ([17])*

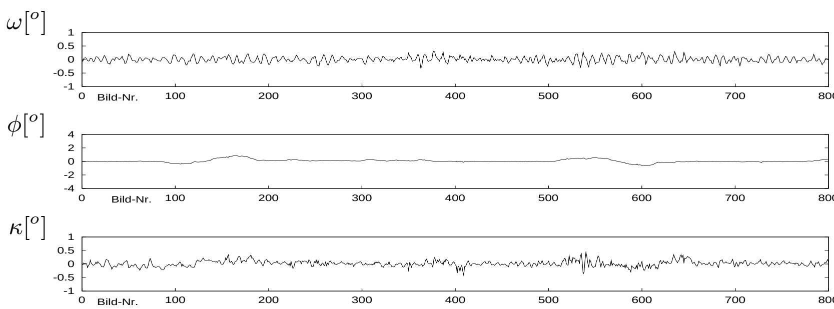

Fig. 3. *Estimated rotation angles of the second image w. r. t the first. Observe the typical vibration in the* ω*.*

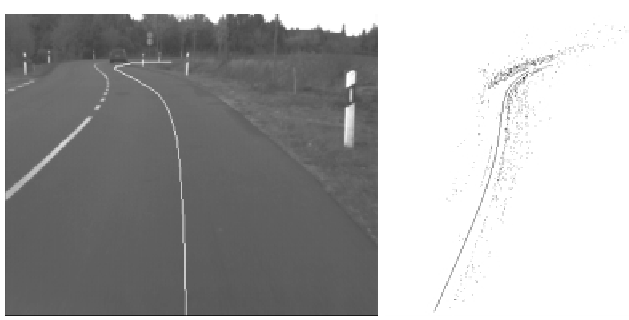

Fig. 4. *Backprojection of trajectory of image sequence with 100 images and 3D point cloud together with trajectory*

• The quality of the motion parameters are much higher when using the image triplet than when only using image pairs (cf. fig. 2).
• The estimation of the rotation angles (cf. fig. 3) reflects the expected behaviour, especially vibrations in nick-angle, i. e. the oscillations of ω around the horizontal x-axis orthogonal to the speed vector.
• The approximate values obtained from the image triplets were sufficiently accurate to guarantee convergence of a global ML-estimation with a bundle adjustment (cf. fig. 4).

Appendix

We give the solution for the optimization problem (4) derived from the Gauss-Helmert-model. For solving this nonlinear problem in an iterative manner we need approximate values $\widehat{\boldsymbol{\beta}}^{(0)}$ and $\widehat{\boldsymbol{y}}^{(0)}$ for the unknowns $\widehat{\boldsymbol{\beta}} = \widehat{\boldsymbol{\beta}}^{(0)} + \widehat{\boldsymbol{\Delta\beta}}$ and $\widehat{\boldsymbol{y}} = \widehat{\boldsymbol{y}}^{(0)} + \widehat{\boldsymbol{\Delta y}}$ which obtain corrections $\widehat{\boldsymbol{\Delta\beta}}$ and $\widehat{\boldsymbol{\Delta y}}$ in an iterative manner. With the Jacobians

$$A = \left(\frac{\partial \boldsymbol{g}(\boldsymbol{\beta},\boldsymbol{y})}{\partial \boldsymbol{\beta}}\right)\Bigg|_{\substack{\beta=\widehat{\beta}^{(0)}\\ y=\widehat{y}^{(0)}}}, B = \left(\frac{\partial \boldsymbol{g}(\boldsymbol{\beta},\boldsymbol{y})}{\partial \boldsymbol{y}}\right)\Bigg|_{\substack{\beta=\widehat{\beta}^{(0)}\\ y=\widehat{y}^{(0)}}}, H = \left(\frac{\partial \boldsymbol{h}(\boldsymbol{\beta})}{\partial \boldsymbol{\beta}}\right)\Bigg|_{\beta=\widehat{\beta}^{(0)}} \tag{23}$$

and the relation $\widehat{\boldsymbol{\Delta y}} = (\boldsymbol{y} - \widehat{\boldsymbol{y}}^{(0)}) - \widehat{\boldsymbol{e}}$ we obtain the linear constraints $\boldsymbol{g}(\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{y}}) = \boldsymbol{g}(\widehat{\boldsymbol{\beta}}^{(0)}, \widehat{\boldsymbol{y}}^{(0)}) + A\,\widehat{\boldsymbol{\Delta\beta}} + B\,\widehat{\boldsymbol{\Delta y}}$ or $\boldsymbol{g}(\widehat{\boldsymbol{\beta}}, \widehat{\boldsymbol{y}}) = \boldsymbol{c}_g + A\,\widehat{\boldsymbol{\Delta\beta}} - B\,\widehat{\boldsymbol{e}}$ and $\boldsymbol{h}(\widehat{\boldsymbol{\beta}}) = \boldsymbol{c}_h + H\,\widehat{\boldsymbol{\Delta\beta}}$ with $\boldsymbol{c}_g = \boldsymbol{g}(\widehat{\boldsymbol{\beta}}^{(0)}, \widehat{\boldsymbol{y}}^{(0)}) + B\,(\boldsymbol{y} - \widehat{\boldsymbol{y}}^{(0)})$ and $\boldsymbol{c}_h = \boldsymbol{h}(\widehat{\boldsymbol{\beta}}^{(0)})$ are the contradictions between the approximate values for the unknown parameters and the given observations and among the approximate values for the unknowns.

Setting the partials of Φ (4) zero yields

$$\frac{\partial \Phi}{\partial \widehat{\boldsymbol{y}}^{\mathsf{T}}} = -Q_{yy}^{-1}\widehat{\boldsymbol{e}} + B^{\mathsf{T}}\boldsymbol{\lambda} = \boldsymbol{0} \qquad \frac{\partial \Phi}{\partial \widehat{\boldsymbol{\beta}}^{\mathsf{T}}} = A^{\mathsf{T}}\boldsymbol{\lambda} + H^{\mathsf{T}}\boldsymbol{\mu} = \boldsymbol{0} \tag{24}$$

$$\frac{\partial \Phi}{\partial \boldsymbol{\lambda}^{\mathsf{T}}} = \boldsymbol{c}_g + A\,\widehat{\boldsymbol{\Delta\beta}} - \boldsymbol{B}\widehat{\boldsymbol{e}} = \boldsymbol{0} \qquad \frac{\partial \Phi}{\partial \boldsymbol{\mu}^{\mathsf{T}}} = \boldsymbol{c}_h + H\,\widehat{\boldsymbol{\Delta\beta}} = \boldsymbol{0} \tag{25}$$

From (24a) follows the relation

$$\widehat{\boldsymbol{e}} = Q_{yy}B^{\mathsf{T}}\boldsymbol{\lambda} \tag{26}$$

When substituting (26) into (25a), solving for $\boldsymbol{\lambda}$ yields

$$\boldsymbol{\lambda} = (B\,Q_{yy}B^{\mathsf{T}})^{-1}(\boldsymbol{c}_g + A\,\widehat{\boldsymbol{\Delta\beta}}) \tag{27}$$

Substitution in (24b) yields the symmetric normal equation system

$$\begin{pmatrix} A^{\mathsf{T}}(B\,Q_{yy}B^{\mathsf{T}})^{-1}A & H^{\mathsf{T}} \\ H & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} \widehat{\boldsymbol{\Delta\beta}} \\ \boldsymbol{\mu} \end{pmatrix} = \begin{pmatrix} -A^{\mathsf{T}}(B\,Q_{yy}B^{\mathsf{T}})^{-1}\boldsymbol{c}_g \\ -\boldsymbol{c}_h \end{pmatrix} \tag{28}$$

The Lagrangian multipliers can be obtained from (27) which then yields the estimated residuals in (26). The estimated variance factor is given by

$$\widehat{\sigma}^2 = \frac{\widehat{\boldsymbol{e}}^{\mathsf{T}}Q_{yy}^{-1}\widehat{\boldsymbol{e}}}{G + H - U} \tag{29}$$

The number R of contraints above the number $U - H$, which is nessessary for determiniming the unknown parameters, the *redundancy* is the denominator $R = G - (U - H)$. We finally obtain the *estimated* covariance matrix

$$\widehat{\boldsymbol{\Sigma}}_{\widehat{\beta}\widehat{\beta}} = \widehat{\sigma}^2 Q_{\widehat{\beta}\widehat{\beta}} \tag{30}$$

of the estimated parameters, where $Q_{\widehat{\beta}\widehat{\beta}}$ results from the inverted reduced normal equation matrix using $\overline{N} = A^{\mathsf{T}}(B\,Q_{yy}B^{\mathsf{T}})^{-1}A$

$$\begin{pmatrix} Q_{\widehat{\beta}\widehat{\beta}} & S^{\mathsf{T}} \\ S & T \end{pmatrix} = \begin{pmatrix} \overline{N} & H^{\mathsf{T}} \\ H & \mathbf{0} \end{pmatrix}^{-1} \tag{31}$$

This expression can be used even if $\overline{N}$ is singular.

References

[1] Abraham, S. (2000): *Kamera-Kalibrierung und Metrische Auswertung von Bildfolgen.* Shaker Verlag, Aachen, 2000.

[2] Faugeras, O.; Mourrain, B. (1995): On the Geometry and Algebra of the Point and Line Correspondencies between N Images. Technical Report 2665, INRIA, 1995.

[3] Faugeras, O.; Papadopoulo, T. (1998): Grassmann-Cayley Algebra for Modeling Systems of Cameras and the Algebraic Equations of the Manifold of Trifocal Tensors. *Trans. of the ROYAL SOCIETY A*, 356:1123–1152, 1998.

[4] Förstner, W.; Gülch, E. (1987): A Fast Operator for Detection and Precise Location of Distinct Points, Corners and Circular Features. In: *Proc. of the Intercomm. Conf. on Fast Processing of Photogr. Data, Interlaken*, pages 281–305, 1987.

[5] Förstner, W. (1999): Choosing Constraints in Image Triplets. Technical report, Institut für Photogrammetrie, Universität Bonn, 1999.

[6] Gauss, C. F. (1809): *Theoria Motus Corpum Coelestium.* Perthes und Besser, Hamburg, 1809.

[7] Haralick, R. M.; Shapiro, L. G. (1993): *Computer and Robot Vision, Vol. II.* Addison-Wesley Publishing Company, 1993.

[8] Hartley, R. (1995): A Linear Method for Reconstruction from Lines and Points. In: *Proc. ICCV*, pages 882–887, 1995.

[9] Helmert, F. R. (1872): *Die Ausgleichungsrechnung nach der Methode der Kleinsten Quadrate.* Teubner, Leipzig, 1872.

[10] Huber, P. J. (1981): *Robust Statistics.* Wiley NY, 1981.

[11] Markoff, A. A. (1912): *Wahrscheinlichkeitsrechnung.* Teubner, Leipzig, 1912.

[12] Mikhail, E. M. (1962): Use of Triplets in Analytical Aerotriangulation. *Photogr. Eng*, 28:625–632, 1962.

[13] Mikhail, E. M. (1963): Use of Two-Directional Triplets in a Sub-Block Approach for Analytical Aerotriangulation. *Photogr. Eng*, 29:1014–1024, 1963.

[14] Mikhail, E. M.; Ackermann, F. (1976): *Observations and Least Squares.* University Press of America, 1976.

[15] Shashua, A.; Anandan, P. (1996): Trilinear Constraints Revisited: Generalized Trilinear Constraints and the Tensor Brightness Constraint. In: *Proc. Image Understanding Workshop*, 1996.

[16] Shashua, A. (1995): Algebraic Functions for Recognition. *IEEE T-PAMI*, 17(8):779–789, 1995.

[17] Steines, B.; Abraham, S. (1999): Metrischer Trifokaltensor für die Auswertung von Bildfolgen. In: Förstner/Buhmann (Eds.), *Mustererkennung '99*, LNCS, 1999.

[18] Steines, B. (1998): Bewegungsschätzung aus Bildfolgen. Diplomarbeit, Institut für Photogrammetrie, Universität Bonn, 1998.

[19] Taubin, G. (1993): An Improved Algorithm for Algebraic Curve and Surface Fitting. In: *Fourth ICCV, Berlin*, pages 658–665, 1993.

[20] Triggs, B. (1998): Optimal Estimation of Matching Constraints. In: *SMILE'98: European Workshop on 3D Structure from Multiple Images of Large-Scale Environments*, Springer, LNCS, pages 658–665, 1998.

[21] Torr, P. H. S.; Zisserman, A. (1997): Robust Parametrization and Computation of the Trifocal Tensor. *Image Vision and Computing*, 15:591–605, 1997.