CALIBRATION OF STATIONARY CAMERAS BY OBSERVING OBJECTS OF EQUAL HEIGHTS ON A GROUND PLANE

Jochen Meidow

Institute of Photogrammetry, University of Bonn
Nußallee 15, D-53115 Bonn, Germany
meidow@ipb.uni-bonn.de

KEY WORDS: Camera, Calibration, Estimation, Measurement, Scene, Accuracy.

ABSTRACT

With the increasing number of cameras the need for plug-and-play calibration procedures arises to realize a subsequent automatic geometric evaluation of observed scenes. An easy calibration procedure is proposed for a non-zooming stationary camera observing objects of initially equal and known heights above a ground plane. The image coordinates of the corresponding foot and head points of these objects serve as observations. For the interior and exterior orientation of the camera a minimal parametrization is introduced with the height of the camera above the ground plane, its pitch and roll angle and the principal distance. With the idea of corresponding foot and head trajectories being homologue, the situation can be reformulated with a virtual second camera observing the scene. Therefore a plane induced homography can be established for the observation model. This special planar homology can be parametrisied with the unknown calibration quantities. Initially the calibration is estimated by observing foot and head points of objects with known heights. In the subsequent evaluation phase the height and positions of unknown objects can be determined. With the same procedure the calibration can be checked and updated if needed. The approach is evaluated with a real scene.

1 INTRODUCTION

Motivation. Metric scene reconstruction is the subject of many vision tasks. With the increasing number of video cameras there is a demand of quick and easy calibration procedures which lower the expenses of camera installations while guaranteeing the desired measurement accuracy. In this paper a calibration procedure is presented for stationary, non-zooming cameras as a contribution to the realization of plug-and-play video cameras. The approach uses the observed foot and head points of object with equal heights on a ground plane. The formulas for the solution of the problem will be assembled and explained and the achievable accuracies for the calibration will be determined as well.

Approach. With a straight line preserving pinhole camera a minimal parametrization is introduced: For the intrinsic camera parameters the principal distance is the crucial parameter which determines the reconstruction. The exterior orientation is realized by the pitch and roll angle of the camera as well as the distance of the projection center to the ground plane (height above ground). For the corresponding foot and head points of imaged objects a so-called *plane induced homography* (Hartley and Zisserman, 2000) can be introduced which maps the foot points into the corresponding head points. Assuming that the head and foot points are identical in the object space, the situation can be reformulated with the help of a second, virtual camera observing the same points. This idea allows to exploit a stereo approach: The motion between both cameras induces a planar homology as a special homography and enables the formulation of constraints between the observations and the unknown parameters. The latter are estimated in a combined adjustment for which in principle no approximation values are needed.

Procedure. The realization of the approach consists of two stages: *(1) Initialization:* Since photogrammetry acquires angles, metric information has to be provided in an initial calibration phase by observing objects of equal and known height. After the collection of sufficient data the initial calibration is performed and then the determination of the height and positions of unknown objects is possible. *(2) Parameter update:* Due to environmental influences the calibration parameters may vary, especially the principal distance, therefore, the parameters have to be checked and updated. By assuming that the camera height above ground is constant, this can be achieved by the observation of possibly other objects of equal but unknown heights.

Notation. We denote vectors of the image and the camera coordinate systems with small boldface letters, e. g. $\boldsymbol{x}$, and coordinates in the object coordinate system with capital boldface letters, e. g. $\boldsymbol{X}$. Vectors and matrices are denoted with slanted letters, matrices sans-serif, thus $\boldsymbol{x}$ or $\boldsymbol{\mathsf{R}}$. Homogeneous vectors and matrices, which represent the same object when multiplied with a scalar $\lambda \neq 0$, are denoted with upright letters, e. g. **x** or K. We use the skew symmetric matrix

$$\boldsymbol{\mathsf{S}}(\boldsymbol{x}) = \begin{pmatrix} 0 & -x_3 & +x_2 \\ +x_3 & 0 & -x_1 \\ -x_2 & +x_1 & 0 \end{pmatrix}$$

of a 3-vector $\boldsymbol{x} = (x_1, x_2, x_3)^\mathsf{T}$ to represent the cross product by $\boldsymbol{a} \times \boldsymbol{b} = \boldsymbol{\mathsf{S}}(\boldsymbol{a})\boldsymbol{b}$. The Euclidean normalization of a vector $\boldsymbol{x}$ is preformed by the operator $\mathsf{N}(\boldsymbol{x}) = \boldsymbol{x}/||\boldsymbol{x}||$.

2 MODELLING

2.1 Parametrisations and Observations

Coordinate Systems. The orientation of the camera in the object coordinate system can be described by the pitch angle α, the roll angle γ and the height Z of the camera (cf. fig. 1). Since the azimuth β of the viewing direction is at our disposal, the rotation matrix from the object to the camera system reads $\boldsymbol{R} = \boldsymbol{R}_Z(\gamma) \cdot \boldsymbol{R}_X(\pi/2 + \alpha)$ and with the normal to the plane in the object coordinate system $\boldsymbol{E} = (0, 0, 1)^\mathsf{T}$ the normal in the camera coordinate systems becomes

$$\boldsymbol{n} = \boldsymbol{R}\boldsymbol{E} = (n_X, n_Y, n_Z)^\mathsf{T}. \tag{1}$$

The relationships between the normal and the angles is

$$\alpha = \arctan\left(\frac{n_Z}{n_Y}\right), \quad \gamma = -\arctan\left(\frac{n_X}{n_Y}\right) \tag{2}$$

and $\boldsymbol{n}^\mathsf{T} = \mathsf{N}\left(\tan(-\gamma), 1, \tan(\alpha)\right)$.

Without loss of generality, the projection center $\boldsymbol{Z} = (0, 0, Z)^\mathsf{T}$ is chosen. The origin of the object coordinate system lies in the reference plane, the Z-axis runs through the projection center of the camera. The Y-axis is defined by the projection of the optical axis onto the plane, the X-axis is perpendicular to both (cf. fig. 1).

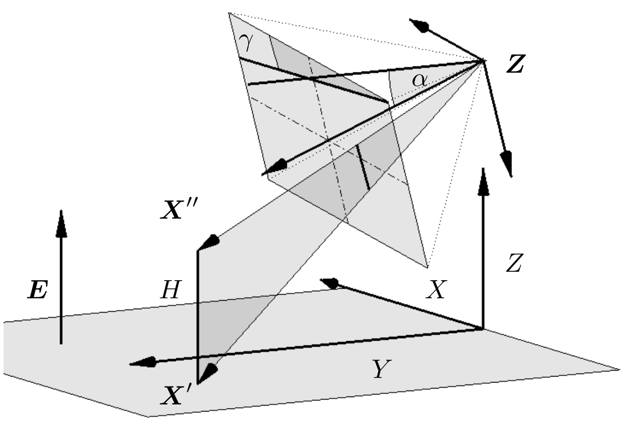

Figure 1: shows the definition of the involved coordinate systems and the projection of a height into the image.

Camera Model. For the camera a straight line preserving pinhole model is introduced with the principal distance c, the scale factor m, the shear s and the principal point (x'_0, y'_0) as the intrinsic camera parameters. With the homogeneous calibration matrix

$$\mathsf{K} = \begin{pmatrix} c & sc & x'_0 \\ 0 & mc & y'_0 \\ 0 & 0 & 1 \end{pmatrix}$$

the homogeneous 3×4–projection matrix $\mathsf{P} = \mathsf{K}\boldsymbol{R}(\boldsymbol{I}_3|-\boldsymbol{Z})$ projects an object point $\boldsymbol{X}_i$ into the image point $\mathbf{x}'_i$ via the linear transformation $\mathbf{x}'_i = \mathsf{P}\mathbf{X}_i$.

With the presented approach and a camera in general position, two of the five intrinsic parameters can be determined — preferably the principal distance and the scale factor. Therefore, initially the used calibration matrix has diagonal shape:

$$\mathsf{K} = \mathrm{Diag}(c, mc, 1).$$

Observations. For each object the four coordinates x'_b, y'_t, y'_b and y'_t (bottom, top) of the foot and head points are available as observations.

2.2 Concept of the Virtual Camera

The mapping of an object foot point $\boldsymbol{x}'_i = (x'_b, y'_b)^\mathsf{T}_i$ into the corresponding head point $\boldsymbol{x}''_i = (x'_t, y'_t)^\mathsf{T}_i$ can be expressed by the projective transformation

$$\mathbf{x}''_i \cong \mathsf{H}\mathbf{x}'_i, \tag{3}$$

called a homography with eight independent parameters due to the homogenity. The 3×3–transformation matrix H is constant for objects of equal height and can be determined by four point correspondences. In the following we show how H can be expressed as a function of the unknown parameters and how a given transformation matrix can be decomposed accordingly.

2.2.1 Virtual Homography. With the notion of corresponding head and foot points being identical in space, the situation can also be described with the help of a second virtual camera (cf. fig. 2). A *plane induced homography* results from the images of two cameras observing the same object on a plane. With the calibration matrices K' and K'' of these two cameras, the distance Z of the first camera to the plane and the baseline vector $\boldsymbol{t}$ the homography reads

$$\mathsf{H} \cong \mathsf{K}''\left(\boldsymbol{R}'' - \frac{1}{Z}\boldsymbol{t}\boldsymbol{n}^\mathsf{T}\right)\mathsf{K}'^{-1}, \tag{4}$$

with the rotation matrix $\boldsymbol{R}''$ of the second camera in respect to the first camera coordinate system; cf. (Faugeras and Lustman, 1988) or (Hartley and Zisserman, 2000) for an alternative derivation. The term in brackets is called the motion matrix $\boldsymbol{M}$. In our case we have one camera observing the scene from two altitudes with unchanged viewing direction, thus $\mathsf{K}' = \mathsf{K}'' = \mathsf{K}$, $\boldsymbol{R}'' = \boldsymbol{I}_3$, and $\boldsymbol{t} = H\boldsymbol{n}$ (cf. fig. 2). The homography (4) becomes

$$\mathsf{H} \cong \mathsf{K}\left(\boldsymbol{I}_3 - \frac{H}{Z}\boldsymbol{n}\boldsymbol{n}^\mathsf{T}\right)\mathsf{K}^{-1}. \tag{5}$$

Observe that the baseline length $\|\boldsymbol{t}\|$ is identical to the object height H.

The transformation (5) is a so-called *planar homology* (Hartley and Zisserman, 2000, p. 585) since with the horizon line $\mathbf{l}' = \mathsf{K}^{-\mathsf{T}}\boldsymbol{n}$ and the vanishing point $\mathbf{v}' = \mathsf{K}\boldsymbol{n}$ — normally the nadir — equation (5) reads

$$\mathsf{H} \cong \boldsymbol{I}_3 + (\mu - 1)\frac{\mathbf{v}'\mathbf{l}'^\mathsf{T}}{\mathbf{v}'^\mathsf{T}\mathbf{l}'}$$

with $(\mu - 1) = H/(Z\mathbf{v}'^\mathsf{T}\mathbf{l}')$. The planar homology has five degrees of freedom — the vertex $\mathbf{v}'$ (2 dof), the axis $\mathbf{l}'$ (2 dof) and the characteristic ratio μ (Semple and Kneebone, 1952) and can therefore be determined by 2.5 point correspondences.

As H contains 5 dof, we can determine two intrinsic parameters in addition to the three parameters α, γ and Z of

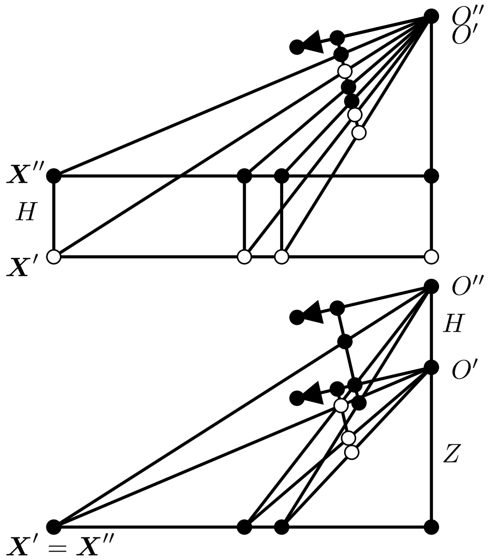

Figure 2: shows the true configuration (top) and the equivalent situation with a second virtual camera (bottom). In both cases the observations of the object foot and head points $\boldsymbol{X}'$ and $\boldsymbol{X}''$ are identical.

the exterior orientation. With the special calibration matrix $\mathsf{K} = \mathrm{Diag}(c, mc, 1)$ the planar homology explicitly reads

$$\mathsf{H} \cong \begin{pmatrix} \frac{Hn_X^2 - Z}{Z} & \frac{Hn_X n_Y}{Zm} & \frac{cHn_X n_Z}{Z} \\ \frac{mHn_X n_Y}{Z} & \frac{Hn_Y^2 - Z}{Z} & \frac{mcHn_Y n_Z}{Z} \\ \frac{Hn_X n_Z}{Zc} & \frac{Hn_Y n_Z}{Zmc} & \frac{Hn_Z^2 - Z}{Z} \end{pmatrix}. \quad (6)$$

Observe that for $m = 1$ the relation $H_{12} \equiv H_{21}$ holds. In many practical cases the roll angle γ equals zero, so that $n_X = 0$ holds and the homography (6) becomes

$$\mathsf{H} \cong \begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{Hn_Y^2 - Z}{Z} & \frac{cmHn_Y n_Z}{Z} \\ 0 & \frac{Hn_Y n_Z}{Zmc} & \frac{Hn_Z^2 - Z}{Z} \end{pmatrix}.$$

as a common specialization. In this case only the principal distance or the scale factor is determinable.

2.2.2 Decomposition of H. Parameter estimation requires approximation values for the unknown calibration parameters. These values can be deduced by a direct estimation of the eight parameters of the common homography (3), if a real-valued decomposition according to (5) is available.

(1) Intrinsic camera parameters. From eq. (6) the principal distance and the scale difference are $c = \sqrt{H_{13}/H_{31}}$ and $m = \sqrt{H_{21}/H_{12}}$,, but for the frequent case of the roll angle $\gamma = 0$ the element H_{31} becomes zero. In this case the principal distance must be computed via $c = \sqrt{H_{23}/H_{32}}/m$ with the known scale factor m.

(2) Exterior orientation. Once the intrinsic parameters c and, where applicable, m have been determined, the motion matrix $\mathsf{M} \cong \mathsf{K}^{-1}\mathsf{H}\mathsf{K}$ can be computed with $\mathsf{K} = \mathrm{Diag}(c, mc, 1)$. The eigenvalue-eigenvector-decomposition of M has three real-valued eigenvectors $\boldsymbol{e}_i$, $i = 1, 2, 3$, with two identical real-valued eigenvalues $\lambda_2 = \lambda_3$ and an individual eigenvalue λ_1. The normal vector of the plane results from the eigenvectors

$$\boldsymbol{n} = \boldsymbol{e}_1 = \mathsf{N}\left(\boldsymbol{e}_2 \times \boldsymbol{e}_3\right)$$

and the ratio of camera and object height from the eigenvalues:

$$\frac{H}{Z} = \frac{\lambda_1 - \lambda_3}{\lambda_1} = \frac{\lambda_1 - \lambda_2}{\lambda_1}.$$

Note that the solution is unambiguous except for a common sign of c and n_Z and the sign of $\boldsymbol{n}$. But the requirement $n_Y > 0$ is reasonable for most camera installations. With the orientation parameters determined in this manner we are able to measure the object height and position.

2.3 3D Object Measurement

Similar formulas for the computation of the height of an object have been developed independently in (Criminisi, 2001) and (Renno et al., 2002, Jones et al., 2002) — on the one hand geometric and on the other hand more algebraic. Below the equivalence of both is shown.

We start from the formulation of the transformation (3) as a condition

$$\boldsymbol{S}(\mathbf{x}_i'')\mathsf{H}\mathbf{x}_i' = \mathbf{0}. \quad (7)$$

With the vertical vanishing point $\mathbf{v}' = \mathsf{K}\boldsymbol{n}$ (the fixed point of the transformation) and the horizon line (fixed line) $\mathbf{l}' = \mathsf{K}^{-\mathsf{T}}\boldsymbol{n}$ (Hartley and Zisserman, 2000) in (5) the condition (7) leads to the formula developed in (Criminisi, 2001)

$$H_i = -\frac{Z}{\mathbf{l}'^{\mathsf{T}}\mathbf{x}_i'} \cdot \frac{\|\boldsymbol{S}(\mathbf{x}_i'')\mathbf{x}_i'\|}{\|\boldsymbol{S}(\mathbf{x}_i'')\mathbf{v}'\|}.$$

by taking the norm of the condition. With the directions $\boldsymbol{m}_i = \mathsf{N}(\mathsf{K}^{-1}\mathbf{x}_i')$ the homography for directions $\boldsymbol{m}_i'' \cong \boldsymbol{M}\boldsymbol{m}_i'$ can be expressed

$$\boldsymbol{S}(\boldsymbol{m}_i'')\boldsymbol{M}\boldsymbol{m}_i' = \mathbf{0}$$

and for the object height the second expression results

$$H_i = -\frac{Z}{\boldsymbol{n}^{\mathsf{T}}\boldsymbol{m}_i'} \cdot \frac{\|\boldsymbol{S}(\boldsymbol{m}_i'')\boldsymbol{m}_i'\|}{\|\boldsymbol{S}(\boldsymbol{m}_i'')\boldsymbol{n}\|}. \quad (8)$$

The position of the object on the plane results from substituting the angular distance $\lambda_i' = -Z/(\boldsymbol{n}^{\mathsf{T}}\boldsymbol{m}_i')$ from the projection center to the foot point $\boldsymbol{X}'$ into the point-direction-form

$$\boldsymbol{X}_i' = (X', Y', Z')_i^{\mathsf{T}} = \boldsymbol{Z} + \lambda_i'\boldsymbol{R}^{\mathsf{T}}\boldsymbol{m}_i' \quad (9)$$

for which $Z_i' = 0$ holds.

The formulas (8) and (9) provide the basis for the object measurement. The calibration procedure is described in the following section.

3 REALISATION

3.1 Calibration Procedure

The proposed calibration procedure consists of two stages: After the initial calibration with objects of equal and known heights, the parameters can be checked and — if needed — updated in the continuous operation phase with new objects of unknown height:

(1) Initial Calibration. After the installation of the camera the foot and head points of the objects have to be measured. Depending on the specific calibration object this can be done manually or with the help of feature extraction. The observed heights may not be arranged on a single straight line in the object space (cf. section 3.3, determinability of the parameters). While the height of the objects has to be known, the height of the camera Z may be introduced as an unknown parameter or — if accessible – as a measured quantity. Approximation values for the unknown parameter can be determined as described below or by a rough guess, e. g. for the roll angle zero is always a good assumption.

(2) Parameter Update. For the continuous operation we assume that the height of the camera does not change, while the other parameters may vary due to environmental influences, for instance temperature. For every new scene t an unknown height H_t is introduced into the adjustment procedure. Since the unknown object heights H_t can vary, relinearisation with few iterations is advisable for every new scene — slightly increasing the computing time. At the same time the measurements yield the position and height of the objects for each image. Furthermore, the adjustment provides the average height for every type of object.

3.2 Approximation Values

Minimizing algebraic distances. The transformation parameters can possibly be determined without the knowledge of approximation values (Hartley and Zisserman, 2000). With the projective transformation $\mathbf{x}''_i \cong \mathsf{H}\mathbf{x}'_i$ written in homogeneous coordinates

$$\begin{pmatrix} u'' \\ v'' \\ w'' \end{pmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \begin{pmatrix} u' \\ v' \\ w' \end{pmatrix}$$

with $\mathbf{x}' = (u', v', w')^\mathsf{T} \cong (x'_b, y'_b, 1)^\mathsf{T}$ and $\mathbf{x}'' = (u'', v'', w'')^\mathsf{T} \cong (x'_t, y'_t, 1)^\mathsf{T}$ we first of all get the constraints between the image coordinates and the homography elements

$$u''_i(gu'_i + hv'_i + i) - w''_i(au'_i + bv'_i + cw'_i) = 0 \quad (10)$$
$$v''_i(gu'_i + hv'_i + i) - w''_i(du'_i + ev'_i + fw'_i) = 0. \quad (11)$$

In compact form $\mathbf{a}_{1i}^\mathsf{T}\mathbf{h} = 0$ and $\mathbf{a}_{2i}^\mathsf{T}\mathbf{h} = 0$ with the 9-vectors

$$\begin{aligned} \mathbf{a}_{1i}^\mathsf{T} &= (-w''_i\mathbf{x}'^\mathsf{T}_i, \quad \mathbf{0}^\mathsf{T}, -u''_i\mathbf{x}'^\mathsf{T}_i) \\ \mathbf{a}_{2i}^\mathsf{T} &= (\quad \mathbf{0}^\mathsf{T}, -w''_i\mathbf{x}'^\mathsf{T}_i, v''_i\mathbf{x}'^\mathsf{T}_i) \end{aligned}$$

and the unknown parameters $\mathbf{h} = (a, b, c, d, e, f, g, h, i)^\mathsf{T}$ we get the homogeneous equation system $\mathsf{A}\mathbf{h} = \mathbf{0}$. The right eigenvector of A for the smallest eigenvalue λ_l is a good estimation for $\mathbf{h}$. With the singular value decomposition $\mathsf{A} = \boldsymbol{UDV}^\mathsf{T}$ the solution is

$$h_k = V_{kl}, \qquad \text{with} \quad k = 1, \ldots, 9 \qquad (12)$$

For numerical reasons a conditioning of the problem is advisable.

Enforcing the homology constraints. The estimation (12) of H does not possess the properties of a planar homology presented in section 2.2.1. Therefore, a least squares adjustment can be done assuming the elements $\boldsymbol{h} = \mathsf{vec}(\mathsf{H})$ as i. i. d. observations. The explicit model of this observation process reads

$$\boldsymbol{h} = \boldsymbol{f}(c, \alpha, \gamma, Z) \quad \text{with} \quad \boldsymbol{\Sigma}_{hh}^{(0)} = \sigma_0^2 \boldsymbol{I}_9 \qquad (13)$$

with the a priori covariance matrix $\boldsymbol{\Sigma}_{hh}^{(0)}$ of the observations and the unknown variance factor σ_0^2. The solution $\widehat{\mathsf{H}}$ minimizes the Frobenius norm $\|\mathsf{H} - \widehat{\mathsf{H}}\|$. Approximation values are taken from the decomposition explained in section 2.2.2.

Although the solution $\widehat{\boldsymbol{h}}$ fulfills the constraints of the planar homology, it is still an approximation since potential individual weights of the observations have not been taken into consideration. Therefore, a subsequent stringent adjustment is necessary.

3.3 Parameter Estimation

Determinability of the Parameters. If the pitch angle α is zero or $90°$ — i. e. the viewing direction is horizontal or towards the nadir — the element n_Z of the normal vector (1) becomes zero. In this case the 2D-homography (6) degenerates to a 1D-homography and the principal distance c is not determinable. If the pitch angle is approximate zero or π, the determination of the parameters is very weak. In this case prior information about the parameters has to be provided. This can easily be done by introducing these values as additional, fictitious observations into the adjustment process explained in the following.

One critical arrangement of the calibrating objects can be observed: if the foot and head points in the image are collinear, the homography degenerates and the parameters are not determinable. Thus not all objects may be situated on a single straight line.

Adjustment Model. For the calibration phases (initial and update) the general non-linear model

$$\boldsymbol{g}(\boldsymbol{l}, \boldsymbol{p}) = \mathbf{0} \qquad \text{with} \quad \boldsymbol{\Sigma}_{ll}^{(0)} = \sigma_0^2 \boldsymbol{P}_{ll}^{-1} \qquad (14)$$

with the constraints between the observations $\boldsymbol{l}$, the parameters $\boldsymbol{p}$ and the a priori covariance matrix of the observations $\boldsymbol{\Sigma}_{ll}^{(0)}$ is arranged, cf. for instance (Mikhail, 1976). The constraints of the model are the eqs. (10) and (11). For technical convenience with $\boldsymbol{p} = (c, \alpha, \gamma, Z, H)^\mathsf{T}$ five parameters have been introduced although just the fraction H/Z is determinable. Depending on the actual calibration phase (initial or update) either H or Z have to be fixed by prior information. Because of the assumption of i. i. d. observation groups the normal equation system for the adjustment model (14) can be built-up sequentially.

To make sure, that the necessary prior information has a constant contribution to the solution, the relative weighting between the observations and the prior information can be controlled by a regularization factor λ. An ad-hoc solution is $\lambda = \mathrm{tr}(\boldsymbol{N})/\mathrm{tr}(\boldsymbol{P}_{pp})$ (Press et al., 1992) with the traces of the normal equation matrix $\boldsymbol{N}$ and the prior weights $\boldsymbol{P}_{pp}$ for the 'observed' parameters. Again, a conditioning of the problem is advisable by a translation and scaling of the image quantities and the principal distance respectively.

Kalman Filter. The sequential build-up of the normal equation system offers the possibility of introducing a discrete Kalman filter (Welch and Bishop, 2002) for the calibration update phase. This is equivalent to a recursive parameter estimation process. To prevent a numerical overflow and the solution to bite, a *memory length* term k can be introduced, which controls the amount of memory used for the actual solution. With $k = 0.9$ for instance, 90 % of the past observations will be used at the present time. After every evaluation step the normal equation matrix, the right-hand-side vector, the sum of squared residuals and the number of conditions have to be updated. The latter becomes real-valued which is as yet practically irrelevant. The parameter k may not affect the unknown object heights H_t as this parameter can vary from scene to scene.

4 EXPERIMENTAL RESULTS

4.1 Observations and Reference Calibration

Observations. For the evaluation of the approach an image of a lecture room was recorded, showing a seating arrangement of chairs of indentical heights (cf. fig. 3). The camera used has an image format of 960×1280 picture elements. The image measurement of the foot points of the chair legs and the top points of the chair backs was done by an operator.

Reference Calibration. For the evaluation of the approach a reference calibration has been carried out for the intrinsic camera parameters as well as for the exterior orientation.

After the recording of the image a calibration field has immediately been captured on location. The intrinsic parameters are then taken from a bundle adjustment. Table 1 summarizes the results of the parameter estimation for the intrinsic parameters.

For the determination of the exterior camera orientation the image points representing the corners of the tables have been measured. Together with the world coordinates of the corresponding points 0.74 m above the ground plane and the interior orientation given in table 1 a spatial resection has been accomplished assuming a standard deviation of 0.02 m for the object coordinates and 2 pel for the images coordinates. From the estimated matrix for the rotation from the object to the camera coordinate system the roll and pitch angle result from (1) and(2). The estimated accuracies result from error propagation and are listed in table 2. The estimated height of the camera above ground has been verified with the help of a measuring tape.

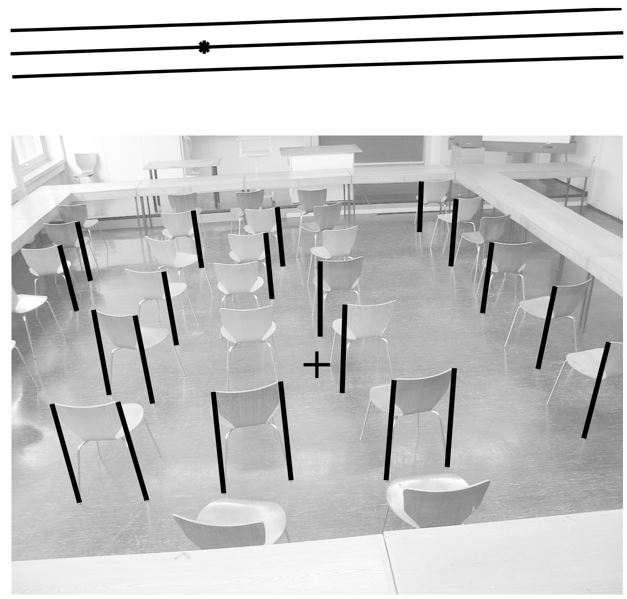

Figure 3: shows the observed corresponding foot and head points as well as the estimated horizon line, its point of gravity and its hyperbolic error band (3σ intervals).

parameter	estimation	estim. std. dev.
principal dist. c	1328.86 pel	2.577 pel
scale factor m	0.9962	$3.377 \cdot 10^{-4}$
principal pt. $\Delta x'_0$	-1.35 pel	1.458 pel
principal pt. $\Delta y'_0$	-4.90 pel	1.389 pel

Table 1: summarizes the results from the intrinsic camera calibration with a test field.

parameter	estimation	estim. std. dev.
pitch angle α	31.2324 deg	0.4479 deg
roll angle γ	0.4847 deg	0.5341 deg
camera position X	3.0611 m	0.0961 m
camera position Y	-2.2095 m	0.0397 m
camera height Z	2.5583 m	0.0830 m

Table 2: summarizes the results of the exterior reference calibration.

4.2 Calibration Results

A height of $H = 0.77$ m have been determined for the chairs in the scene. The results of the direct solution (12) and of the constrained advancement with (13) are summarized in table 3.

For the following calibrations prior information has to be used in order to introduce metric information. For the

parameter	direct sol.	constrained
principal dist. c	1157.8 pel pel	1160.5 pel
pitch angle α	+29.99 deg	+30.01 deg
roll angle γ	-2.52 deg	+0.08 deg
camera height Z	2.49 m	2.49 m

Table 3: shows the results of the direct solution and its constrained add-on.

height of the chairs $H = 0.77$ m, $\sigma_H = 0.02$ m has been introduced. Table 4 summarizes the results of the initial calibration with a redundancy of 38. The process converged after four iterations. The estimated factor $\widehat{\sigma}_0 = 3.75$ lies in the expected magnitude for the precisions of the image points. Figure 3 shows the results qualitative. Drawn in the image is the estimated horizon line with its hyperbolic error band. The position and orientation of the horizon line can be easily checked by visual inspection of the vanishing lines.

parameter	estim.	est. std. dev.
principal dist. c	1196.8 pel	32.3 pel
pitch angle α	+29.37 deg	0.48 deg
roll angle γ	-1.96 deg	0.37 deg
camera height Z	2.53 m	0.05 m

Table 4: shows the results of the initial calibration.

4.3 Object Measurement

The observed and measured chair legs are illustrated in fig. 4 in an upright projection, together with the projection center, the footprint of the principal point and the projection of an image raster. The positions and heights of new, unknown objects can be determined by (8) and (9).

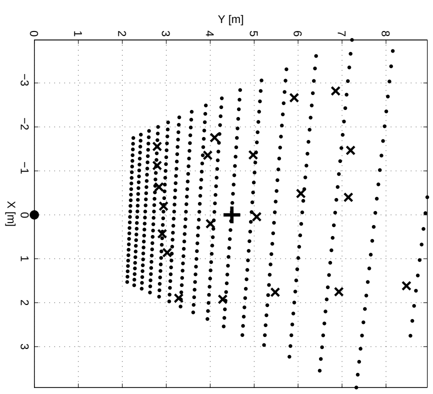

Figure 4: shows the footprints of a image raster and the positions ($\times$) of the chair legs on the ground plane.

5 CONCLUSIONS AND OUTLOOK

Conclusions. An easy camera calibration procedure has been presented for the observation of objects of equal heights on a ground plane. The procedure uses a minimal parametrization for the camera itself and its exterior orientation. Few efforts are associated with the installation; the foot and head points of the objects serve as observations. After an initialization phase with a first scene the approach allows within the continuous operation a parameter check and update if necessary. For the calibration results the single parameter values are less important than the specific parameter combination; the change of one parameter can to some degree be compensated by the others. Due to the sequential build-up of the normal equations, the demand of storage space is minimal. For the set-up of the camera system a pitch angle $> 20°$ and a large aperture angle (or small principal distance) are advisable. Otherwise prior information has be be introduced to cope with the weak geometric configuration. The prior information guarantees but also dominates the solution. The height of the camera should be measured wherever possible in order to impose more geometric constraints onto the solution.

Outlook. In order to eliminate the influence of gross observational errors, a robust estimation is desirable. Furthermore, the integration of other easily available measurements — such as distances in the object space — is advantageous, depending on the precise location to be recorded.

REFERENCES

Criminisi, A., 2001. Accurate Visual Metrology from Single and Multiple Uncalibrated Images. Distinguished Dissertations, Springer, London, Berlin, Heidelberg.

Faugeras, O. and Lustman, F., 1988. Motion and Structure from Motion in a piecewise planar Environment. International Journal of Pattern Recognition in Artificial Intelligence 2, pp. 485–508.

Hartley, R. and Zisserman, A., 2000. Multiple View Geometry in Computer Vision. Cambridge University Press, Cambridge.

Jones, G. A., Renno, J. and Remagnino, P., 2002. Auto-Calibration in Multiple-Camera Surveillance Environments. In: 3rd IEEE Workshop on Performance Evaluation of Tracking and Surveillance (PETS 02), Copenhagen, pp. 40–47.

Mikhail, E. M., 1976. Observations and Least Squares. With Contributions by F. Ackerman. University Press of America, Lanham.

Press, W. H., Teukolsky, S. A., Vetterling, W. T. and Flannery, B. P., 1992. Numerical Recipies in C. Cambridge University Press.

Renno, J., Orwell, J. and Jones, G., 2002. Learning Surveillance Tracking Models for the Self-Calibrated Ground Plane. In: British Machine Vision Conference. Poster Session.

Semple, J. G. and Kneebone, G. T., 1952. Algebraic Projective Geometry. Oxford Univ. Press, New York.

Welch, G. and Bishop, G., 2002. An Introduction to the Kalman Filter. Technical Report TR 95-041, Department of Computer Science, Univ. of North Carolina at Chapel Hill.