University of Passau
Technical Report

Additional constraints for Zhang's closed form solution of the camera calibration problem

Tobias Hanning and René Schöne *

September 13, 2007

Abstract

A known planar model and its observation in the image of a pinhole camera are related by a planar homography. This planar homography is up to a scale factor given by the projection matrix and a rotation and translation. From a number of observations of the same model one can derive the projection matrix by two constraints introduced by the rotation part. These constraints can be transformed to a homogeneous system of linear equations. An additional constraint is needed to avoid the trivial solution. The canonic way is to determine the solution on the unit sphere. But this side condition does not reflect any correlation between the parameters. A solution of this problem may not represent a valid projection matrix. In this article we search for more meaningful restrictions for the solution. For some special but relevant cases we are able to express the problem by a quadratic constraint, which leads to a generalized Eigenvalue problem. We show the profit of the additional constraints in several experimental results.

*{hanning,schoene}@forwiss.uni-passau.de

Contents

1 Introduction

A calibrated camera is the key to every metric 3D reconstruction. The camera mapping describes the process of mapping a point $p \in \mathbb{R}^3$ w. r. t. a reference coordinate system to an image point $i_p \in \mathbb{R}^2$ observed by the camera.

It is defined by an isometric transformation followed by a central projection followed by a reparametrization. The determination of all parameters defining this mapping is a non-linear optimization problem. Many approaches to this problem can be found (see e. g. [Tsa87, WCH92, WM94, HS97] and many more). However, the result of a non-linear optimization depends on a good starting value. We call the task to determine a closed form for the starting value the *camera calibration problem*.

Zhang's approach to the camera calibration problem is based on the observation that the mapping of planar model to its observation can be described by a planar homography which is related to the camera mapping by a scalar ([Zha98]). So, the isometric transformation part of the camera mapping introduces two conditions from which properties of the camera mapping can be derived. These properties can be expressed linear in a six-dimensional vector, whose entries define the reparametrization. The solution of this system of linear equations on the unit sphere represents a solution of the camera calibration problem w. r. t. an algebraic error. But the entries of the solution vector are correlated by the properties from which this vector is derived. This correlation is supressed by the linearization of the problem. Therefore, a minimum w. r. t. the algebraic error does not ensure a valid reparametrization. In particular, an invalid solution is likely for cameras with a low resolution. Nowadays, range cameras are a relevant example of image sensors with a very limited resolution. A range camera provides an image, in which each pixel delivers the distance from the sensor surface to the observed scene. In this article we present some results for a camera based on the Siemens 64×8 Pixel time-of-flight array sensor as presented by Mengel et al. [MLK+07], for a common web camera, and for a camera with a wide angle lens.

In the following we investigate the correlations between the entries of the solution vector. For some restrictions of the camera mapping we are able to present additional constraints for the system of linear equations which are necessary for a valid solution. The main result will be a quadratic constraint for a projection matrix with a known aspect ratio. For a projection matrix with unknown aspect ratio we present a necessary condition for the solution which can also be expressed by a quadratic form. Additionally, we present closed form solutions for the constrained systems of linear equations.

In the next section we define the camera mapping in Cartesian and in homogeneous coordinates. The representation in homogeneous coordinates allows to code the projection and reparametrization of a camera mapping by the so called projection matrix. Following Zhang ([Zha98]) we formulate two conditions for this projection matrix derived by observing planar patterns in the third section . These conditions are necessary for the parameters of the projection matrix. In the fourth section a linear guess of these parameters applying the conditions is derived. To ensure a valid solution of this linear guess, we propose additional constraints for the starting point for several restricted camera models. We complete this presentation by some experimental results.

2 The camera mapping

2.1 Camera parametrization in a Cartesian coordinate system

The first step of every camera mapping is an isometric coordinate transformation from the reference coordinate system to the camera coordinate system. An isometric transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ can be described by a rotation $R \in \mathbb{R}^{3\times 3}$ and a translation $t \in \mathbb{R}^3$ as $T(p) = Rp + t$ for $p \in \mathbb{R}^3$. All parameters describing this transformation are called *extrinsic parameters*.

The next step is the projection of 3D point on the image plane. Let Π denote the central projection w. r. t. the z-coordinate:

$$\Pi : \begin{array}{l} \mathbb{R}^2 \times \mathbb{R} \setminus \{0\} \to \mathbb{R}^2 \\ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x/z \\ y/z \end{pmatrix} \end{array} . \tag{1}$$

Note that Π describes no change of units. The result of the central projection describes a point w. r. t. the camera coordinate system. The last step accomplishes the change of the camera coordinate system to the image coordinate system. We set

$$P : \begin{array}{l} \mathbb{R}^2 \to \mathbb{R}^2 \\ \begin{pmatrix} u \\ v \end{pmatrix} \mapsto \begin{pmatrix} \alpha & \gamma \\ 0 & \beta \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} u_0 \\ v_0 \end{pmatrix} \end{array} \tag{2}$$

where (u_0, v_0) denotes the intersection of the optical axis with the image plane w. r. t. the image coordinate system and is called *principal point*. The parameters α and β are scale factors in the two directions of the imaging sensor. If f is the focal length of the camera and $d_u \times d_v$ is the size of a CCD-element, then the parameters α and β can be interpreted as $\alpha = \frac{f}{d_u}$ and $\beta = \frac{f}{d_v}$. γ describes the skewness between the axes of the pixel coordinate system. For $\gamma = 0$ the coordinate axes of the image coordinate system are perpendicular. For real sensors the image coordinate system should be orthogonal, but due to a misaligned lens it may appear skewed in the projection matrix.

After all, our camera mapping $\mathcal{K} : \mathbb{R}^3 \to \mathbb{R}^2$ can be parametrized by $\mathcal{K} = P \circ \Pi \circ T$. However, real camera mappings often show a different behavior. A distortion mapping $\delta : \mathbb{R}^2 \to \mathbb{R}^2$ in the image plane w. r. t. the camera coordinate system must be defined to model the camera mapping for real cameras. The distortion is placed after the projection and before the transformation from the camera coordinate system to the image coordinate system: $\mathcal{K} = P \circ \delta \circ \Pi \circ T$. The most common distortion model is the one of a radial distortion (see e.g. [Atk96]).

A camera model with no distortion ($\delta = \mathrm{id}$) is called *pure pinhole camera model.* All parameters which describe the mapping $P \circ \delta$ are called *intrinsic camera parameters.*

2.2 Camera parametrization in homogeneous coordinates

For a vector $x = (x_1, \ldots, x_n)^t \in \mathbb{R}^n$ let $\tilde{x} = (x_1, \ldots, x_n, 1)^t$ be the canonic embedding of x into the projective completion $\mathbb{P}(\mathbb{R}^n)$ of $\mathbb{R}^n$. $\tilde{x}$ are the homogeneous coordinates of x. The transition from a Cartesian reference coordinate system to homogeneous coordinates allows us to denote coordinate system transformations as matrices. Since a camera mapping of a pure pinhole camera model consists of coordinate system transformations the mapping of a pure pinhole camera can be described by a matrix. We define

$$\tilde{T} := \begin{bmatrix} Rt \end{bmatrix} \in \mathbb{R}^{3\times 4}, \tag{3}$$

where the first three columns of $\tilde{T}$ are defined by the columns of a rotation matrix R and the last column represents a translation t. Furthermore, we call

$$\tilde{P} := \begin{pmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{pmatrix} \tag{4}$$

the *projection matrix.* Assuming a pure pinhole camera without distortion

$$\tilde{\mathcal{K}} := \tilde{P}\tilde{T} \in \mathbb{R}^{3\times 4} \tag{5}$$

describes the camera mapping. The projection $(u, v) \in \mathbb{R}^2$ of a point $(x, y, z) \in \mathbb{R}^3$ by the camera mapping fulfills

$$\lambda \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \tilde{\mathcal{K}} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \tag{6}$$

for a scalar $\lambda \in \mathbb{R}$. Equation (6) is called *pinhole model equation.* So, for every point $p = (x, y, z)^t$ the projection (u, v) by the camera $\mathcal{K}$ can be obtained from $(u', v', w') = \tilde{\mathcal{K}}(x, y, z, 1)^t$ by setting $(u, v) = (\frac{u'}{w'}, \frac{u'}{w'})^t$, if $w' \neq 0$. Otherwise, the point has no image point and accordingly is a point at infinity in the projective sense.

3 Camera calibration

The process of determining the extrinsic and intrinsic camera parameters is called *camera calibration.* Most algorithms determine these parameters by observing a well known model $\mathbf{M}$, where $\mathbf{M} \subset \mathbb{R}^3$ is a finite set of points w. r. t. the reference coordinate system.

For every $p \in \mathbf{M}$ we denote $i_p \in I \subset \mathbb{R}^2$ for the observed projection of p in the image plane I w. r. t. the image coordinate system.

The standard camera calibration minimizes the function

$$\Phi : \mathcal{K} \mapsto \sum_{p \in \mathbf{M}} \|i_p - \mathcal{K}(p)\|^2 \ . \tag{7}$$

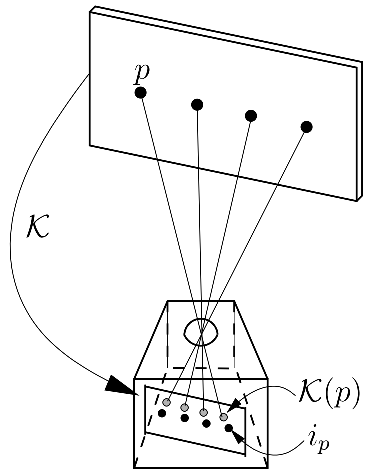

Figure 1: Minimized distance of the calibration

This means that the calibration minimizes the Euclidean distances of the observed image points i_p to the model points projected on the image plane by the camera mapping $\mathcal{K}$ (see Fig. 1).

For flexible calibration purposes the set of points $\mathbf{M}$ is often defined by a grid of points on a plane (e. g. obtained by processing images of an observed checkerboard pattern). Obviously, more than one observation of such a planar calibration pattern is needed to yield a sufficient result. A planar pattern provides great flexibility when using several observations at different positions and angles. With multiple targets the number of parameters in the error function Φ arises: one has to determine the position of every pattern. For N positions of the model the error function becomes

$$\Phi_N : (P, \delta, T_1, \ldots, T_N) \mapsto \sum_{n=1}^{N} \sum_{p \in \mathbf{M}} \|i_{n,p} - P \circ \delta \circ \Pi \circ T_n(p)\|^2 \tag{8}$$

where $i_{n,p}$ is the observation of $p \in \mathbf{M}$ in the n-th image.

Obviously, the minimization of Φ_N defines a non-linear problem. Such a problem depends crucially on an adequate starting value. In the next section we follow the ideas of Zhang in [Zha98] to obtain a closed form solution for the intrinsic camera parameter in an algebraic sense.

One should note that the knowledge of the intrinsic parameters allows an estimation of the extrinsic parameters (see e. g. [HSG06]).

4 A starting solution from constraints of the observed homographies

Assuming a pure pinhole camera an image of an calibration object defines a homography from the object's coordinate system to the image coordinate system. For a fixed observation $(i_p)_{p\in\mathbf{M}}$ of a model $\mathbf{M}$ the homography $H \in \mathbb{R}^{3\times 4}$ should minimize

$$\sum_{p\in\mathbf{M}} \|i_p - \Pi(H\tilde{p})\|^2. \tag{9}$$

With the pinhole equation (6) for an ideal observation of the model $\mathbf{M}$

$$\lambda H = \tilde{P}\tilde{T} = \tilde{P}\left[Rt\right] \tag{10}$$

must hold, for a pure pinhole camera $\tilde{P}\tilde{T}$ and a $\lambda \in \mathbb{R}$

Let $H = (h_1\ h_2\ h_3\ h_4) \in \mathbb{R}^{3\times 4}$ with columns $h_i \in \mathbb{R}^3$ for $i = 1, \ldots, 4$ be the homography which minimizes (9). A planar model $\mathbf{M}$ defines a reference coordinate system such that every z-coordinate of a point in the model plane is zero. This means it is

$$H \begin{pmatrix} x \\ y \\ 0 \\ 1 \end{pmatrix} = (h_1\ h_2\ h_4) \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{11}$$

for every point $(x, y, 0)^t$ within the model plane. Therefore, the column h_3 is not of interest when we observe a planar pattern. The determination of the planar homography $H' = (h_1\ h_2\ h_4)$ defines a non-linear problem with 8 degrees of freedom, since H' is determined up to scale factor. A linear estimation of a starting point can be found in [Zha98]. A normalization of the input data as described in [Har97] will improve the initial guess.

Let now $r_1, r_2, r_3 \in \mathbb{R}^3$ be the columns of a rotation matrix $R = (r_1\ r_2\ r_3)$. For a planar target we derive by Equation 6

$$\lambda \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \tilde{P}\tilde{T} \begin{pmatrix} x \\ y \\ 0 \\ 1 \end{pmatrix} = \tilde{P}(r_1\ r_2\ t) \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{12}$$

for a scalar $\lambda \in \mathbb{R}$. Therefore, it follows $\lambda(h_1\ h_2\ h_4) = \tilde{P}(r_1\ r_2\ t)$ with Equation 10 and Equation 11.

The columns r_1 and r_2 of the rotation matrix R are orthonormal. With $\lambda\tilde{P}^{-1}(h_1\ h_2\ h_4) = \begin{bmatrix} r_1 & r_2 & t \end{bmatrix}$ we obtain

$$0 = r_1^t r_2 = \lambda^2 \left(\tilde{P}^{-1}h_1\right)^t \tilde{P}^{-1}h_2 = \lambda^2 h_1^t (\tilde{P}^{-1})^t \tilde{P}^{-1} h_2 \ . \tag{13}$$

Hence, for $\lambda \neq 0$

$$h_1^t \tilde{P}^{-t} \tilde{P}^{-1} h_2 = 0 \tag{14}$$

holds, where $\tilde{P}^{-t}$ denotes $(\tilde{P}^t)^{-1}$. Analogously, with $r_1^t r_1 = r_2^t r_2 = 1$ we obtain

$$h_1^t \tilde{P}^{-t} \tilde{P}^{-1} h_1 - h_2^t \tilde{P}^{-t} \tilde{P}^{-1} h_2 = 0 \ . \tag{15}$$

Equation 14 and 15 define two constraints for the projection matrix $\tilde{P}$. These equations can also be derived by analyzing the image of the absolute conic $\tilde{P}^{-t}\tilde{P}^{-1}$ and the circular points. For a better readability we omit the details of this approach. See [HZ00] for an elaborated introduction.

4.1 Zhang's starting solution

We now present the original approach of Zhang as presented in [Zha98] to the camera calibration problem: The matrix $B = \tilde{P}^{-t}\tilde{P}^{-1}$ is given by

$$B = \begin{pmatrix} \frac{1}{\alpha^2} & -\frac{\gamma}{\alpha^2\beta} & \frac{v_0\gamma - u_0\beta}{\alpha^2\beta} \\ -\frac{\gamma}{\alpha^2\beta} & \frac{\gamma^2+\alpha^2}{\alpha^2\beta^2} & -\frac{\gamma(v_0\gamma-u_0\beta)+v_0\alpha^2}{\alpha^2\beta^2} \\ \frac{v_0\gamma-u_0\beta}{\alpha^2\beta} & -\frac{\gamma(v_0\gamma-u_0\beta)+v_0\alpha^2}{\alpha^2\beta^2} & \frac{(v_0\gamma-u_0\beta)^2+v_0^2\alpha^2}{\alpha^2\beta^2}+1 \end{pmatrix} . \tag{16}$$

Since B is symmetric, it can be parametrized by six values: $b = (b_1, b_2, b_3, b_4, b_5, b_6)^t$ with $b_1 := B_{11}, b_2 := B_{12}, b_3 := B_{22}, b_4 := B_{13}, b_5 := B_{23}$, and $b_6 := B_{33}$. Let $h_1 := (h_{11}, h_{21}, h_{31})^t$ and $h_2 := (h_{12}, h_{22}, h_{32})^t$ be the first and the second column of H'. Then it is

$$\begin{aligned} h_1^t B h_2 &= (h_{11}h_{12}, h_{12}h_{21} + h_{22}h_{11}, h_{22}h_{21}, \\ &\qquad h_{12}h_{31} + h_{32}h_{11}, h_{22}h_{31} + h_{32}h_{21}, h_{32}h_{31})b, && (17) \\ h_1^t B h_1 &= (h_{11}^2, 2h_{11}h_{21}, h_{21}^2, 2h_{11}h_{31}, 2h_{21}h_{31}, h_{31}^2)b, && (18) \\ \text{and } h_2^t B h_2 &= (h_{12}^2, 2h_{12}h_{22}, h_{22}^2, 2h_{12}h_{32}, 2h_{22}h_{32}, h_{32}^2)b. && (19) \end{aligned}$$

Therefore, Equation 14 and Equation 15 are linear in b. Hence,

$$\begin{pmatrix} h_{11}h_{12} & h_{12}h_{21}+h_{22}h_{11} & h_{22}h_{21} & h_{12}h_{31}+h_{32}h_{11} & h_{22}h_{31}+h_{32}h_{21} & h_{32}h_{31} \\ h_{11}^2-h_{12}^2 & 2h_{11}h_{21}-2h_{12}h_{22} & h_{21}^2-h_{22}^2 & 2h_{11}h_{31}-2h_{12}h_{32} & 2h_{21}h_{31}-2h_{22}h_{32} & h_{31}^2-h_{32}^2 \end{pmatrix} b = 0 \tag{20}$$

must hold for every observed homography H. For n observations we get n homographies and hence $2n$ equations which we stack in a matrix V, such that $Vb = 0$ must hold.

To avoid the trivial solution one can compute $b^* = \min_{\|b\|=1} \|Vb\|$, which is the Eigenvector associated to the smallest eigenvalue of V^tV. b^* encodes B up to a scale factor λ. From $b^* = \lambda(B_{11}, B_{12}, B_{22}, B_{13}, B_{23}, B_{33})^t$ Zhang derives

$$\begin{aligned} v_0 &= \frac{b_2^* b_4^* - b_1^* b_5^*}{b_1^* b_3^* - {b_2^*}^2}, & \beta &= \sqrt{\frac{\lambda b_1^*}{b_1^* b_3^* - {b^*}_2^2}}, \\ \lambda &= b_6^* - \frac{{b_4^*}^2 + v_0(b_2^* b_4^* - b_1^* b_5^*)}{b_1^*}, & \gamma &= -\frac{b_2^* \alpha^2 \beta}{\lambda}, \\ \alpha &= \sqrt{\frac{\lambda}{b_1^*}}, \text{ and} & u_0 &= \frac{\gamma v_0}{\beta} - \frac{b_4^* \alpha^2}{\lambda}. \end{aligned} \tag{21}$$

It should be noticed that the constraint $\|b^*\| = 1$ does not guarantee that all assignments in (21) are well defined.

Another way to obtain the projection matrix $\tilde{P}$ from b^* is given by the Cholesky factorization of the matrix $\begin{pmatrix} b_1^* & b_2^* & b_4^* \\ b_2^* & b_3^* & b_5^* \\ b_4^* & b_5^* & b_6^* \end{pmatrix}$ defined by b^* (see e. g. [HZ00]). For a positive definite symmetric matrix A the Cholesky factorization $A = C^tC$ ensures that the diagonal entries in the upper triangular matrix C are positive. The scale factor λ can be obtained simply be the fact that $\tilde{P}_{33} = 1$ holds for the projection matrix. However, this requires that b^* encodes a positive definite matrix.

4.2 A starting solution with known center and zero skew

To simplify the problem we assume that the principal point w. r. t. the image coordinate system coincides with the center of the image. In this case the parameters u_0 and v_0 of the projection matrix are known. Furthermore, we assume that the imaging device is located strictly perpendicular to the optical axis of the lens and that the axes of the imaging device define an orthogonal coordinate system. This yields $\gamma = 0$ in the projection matrix. In this case we are able to parametrize B by using two variables

$$B = \begin{pmatrix} \frac{1}{\alpha^2} & 0 & -\frac{u_0}{\alpha^2} \\ 0 & \frac{1}{\beta^2} & -\frac{v_0}{\beta^2} \\ -\frac{u_0}{\alpha^2} & -\frac{v_0}{\beta^2} & \frac{u_0}{\alpha^2} + \frac{v_0^2}{\beta^2} + 1 \end{pmatrix} = \begin{pmatrix} b_1 & 0 & -u_0 b_1 \\ 0 & b_2 & -v_0 b_2 \\ -u_0 b_1 & -v_0 b_2 & u_0^2 b_1 + v_0^2 b_2 + 1 \end{pmatrix} \tag{22}$$

with $b_1 = \frac{1}{\alpha^2}, b_2 = \frac{1}{\beta^2}$. This means that in this case the parameters b_1, b_2 are uncorrelated. The constraints (14) and (15) lead to a system of linear equations, since for $b = (b_1, b_2)^t$

it is

$$
\begin{aligned}
(23)\qquad h_1^t B h_2 &= \big(h_{11}h_{12} - u_0(h_{11}h_{32} + h_{31}h_{12}) + u_0^2 h_{31}h_{32},\\
&\quad h_{21}h_{22} - v_0(h_{21}h_{32} + h_{31}h_{22}) + v_0^2 h_{31}h_{32}\big)b\\
&\quad + h_{32}h_{31}\\
(24)\qquad h_1^t B h_1 &= \big(h_{11}^2 - 2u_0h_{11}h_{31} + u_0^2h_{31}^2, h_{21}^2 - 2v_0h_{21}h_{31} + v_0^2h_{31}^2)\big)b + h_{31}^2\\
(25)\qquad h_2^t B h_2 &= \big(h_{12}^2 - 2u_0h_{12}h_{32} + u_0^2h_{32}^2, h_{22}^2 - 2v_0h_{22}h_{32} + v_0^2h_{32}^2)\big)b + h_{32}^2.
\end{aligned}
$$

So, the constraints $h_1^t Bh_2 = 0$ and $h_1^t Bh_1 - h_2^t Bh_2 = 0$ yield a system of linear equations. Here, a solution can be obtained by one observation. Otherwise, for $n > 1$ observations we obtain the linear least squares problem $\|V'b - h'\|$. In fact, the closed form solution implemented in the camera calibration routine of Intel's Open Source Computer Vision library (OpenCV, [Bra02]) is based on this method. To be more precise the OpenCV uses the operation $H' := \left(\begin{smallmatrix} 1 & 0 & -u_0 \\ 0 & 1 & -v_0 \\ 0 & 0 & 1 \end{smallmatrix}\right) H$ on every observed homography H to obtain a modified problem with $(u_0, v_0) = (0, 0)$. In the OpenCV the resulting least squares problem is solved by the pseudo-inverse. However, a householder transformation will provide a more numerical stable solution of $\|V'b - h'\|$. Furthermore, a linear constraint can be formulated: For a valid solution b^* the value $(u_0^2\ v_0^2)b^*$ should be positive. Since b^* defines the parameters of the projection matrix only up to scale factor, we can also demand that $(u_0^2\ v_0^2)b^* = 1$ should hold. In addition to that we can demand that b_1 and b_2 have the same algebraic sign, thus $b_1b_2 > 0$ should hold. In this case a valid solution can be constrained to fulfill $b_1b_2 = 1$. We do not carry out this case, because the solution is obvious.

4.3 A starting solution with known center and unknown skew

A skew in the image coordinate system introduces a mixed term of u_0 and v_0 in B_{33}. For a known optical center (u_0, v_0) w. r. t. the image coordinate system we achieve

$$
(26)\quad B = \begin{pmatrix}
\frac{1}{\alpha^2} & -\frac{\gamma}{\alpha^2\beta} & -u_0\frac{1}{\alpha^2} + v_0\frac{\gamma}{\alpha^2\beta} \\
-\frac{\gamma}{\alpha^2\beta} & \frac{\gamma^2}{\alpha^2\beta^2} + \frac{1}{\beta^2} & u_0\frac{\gamma}{\alpha^2\beta} - v_0\left(\frac{\gamma^2}{\alpha^2\beta^2} + \frac{1}{\beta^2}\right) \\
-u_0\frac{1}{\alpha^2} + v_0\frac{\gamma}{\alpha^2\beta} & u_0\frac{\gamma}{\alpha^2\beta} - v_0\left(\frac{\gamma^2}{\alpha^2\beta^2} + \frac{1}{\beta^2}\right) & u_0^2\frac{1}{\alpha^2} + v_0^2\left(\frac{\gamma^2}{\alpha^2\beta^2} + \frac{1}{\beta^2}\right) - u_0v_0\frac{2\gamma}{\alpha^2\beta} + 1
\end{pmatrix}.
$$

Hence, $b = (b_1, b_2, b_3, b_4)^t$ with $b_1 = \frac{1}{\alpha^2}, b_2 = \frac{1}{\beta^2}, b_3 = \frac{\gamma}{\alpha^2\beta}, b_4 = \frac{\gamma^2}{\alpha^2\beta^2}$ yields

$$
(27)\qquad B = \begin{pmatrix}
b_1 & -b_3 & -u_0b_1 + v_0b_3 \\
-b_3 & b_4 + b_2 & u_0b_3 - v_0(b_4 + b_2) \\
-u_0b_1 + v_0b_3 & u_0b_3 - v_0(b_4 + b_2) & u_0^2b_1 + v_0^2(b_4 + b_2) - 2u_0v_0b_3 + 1
\end{pmatrix}.
$$

As in the case without skew, Equation 14 and 15 lead to a system of linear equations. Notice that at least two observations are needed to obtain a unique solution. Furthermore, the entries of the vector b are correlated: For example it is $\frac{b_3^2}{b_1} = b_4$. One should keep in mind that a known center is a strong assumption. In practical applications the error caused by a fixed but wrong location of the center is much greater than the one by leaving out a parameter for the skewness of the image axes. In particular, considering that this parameter should be very small, we omit a detailed analysis of this case.

4.4 A starting solution with known aspect ratio and no skew

An assumption, which is very common for cameras with low resolution, is that the pixels are actually squared areas. But special camera sensors may also have a very unusual aspect ratio: In particular, for the sensor of the Siemens range camera ([MLK+07]) the pixel size is $130\mu\text{m} \times 300\mu\text{m}$.

In this section we assume that we know the size of sensor element. This means that we know the factor $c \in \mathbb{R}_+$ such that it is $\beta = c\alpha$ in the projection matrix. Furthermore, we assume $\gamma = 0$ in the projection matrix. Therefore, we achieve the restricted projection matrix

(28)
$$\tilde{P} = \begin{pmatrix} \alpha & 0 & u_0 \\ 0 & c\alpha & v_0 \\ 0 & 0 & 1 \end{pmatrix}.$$

In this case it is
(29)
$$B = \tilde{P}^{-t}\tilde{P}^{-1} = \begin{pmatrix} \frac{1}{\alpha^2} & 0 & -\frac{u_0}{\alpha^2} \\ 0 & \frac{1}{c^2\alpha^2} & -\frac{v_0}{c^2\alpha^2} \\ -\frac{u_0}{\alpha^2} & -\frac{v_0}{c^2\alpha^2} & \frac{u_0^2}{\alpha^2} + \frac{v_0^2}{c^2\alpha^2} + 1 \end{pmatrix} = \frac{1}{\alpha} \begin{pmatrix} \frac{1}{\alpha} & 0 & -\frac{u_0}{\alpha} \\ 0 & \frac{1}{c^2\alpha} & -\frac{v_0}{c^2\alpha} \\ -\frac{u_0}{\alpha} & -\frac{v_0}{c^2\alpha} & \frac{u_0^2}{\alpha} + \frac{v_0^2}{c^2\alpha} + \alpha \end{pmatrix}.$$

B can be described by $b = (b_1, b_2, b_3, b_4)^t$ with $b_1 := \frac{1}{\alpha}, b_2 := -\frac{u_0}{\alpha}, b_3 := -\frac{v_0}{\alpha}$, and $b_4 := \frac{u_0^2}{\alpha} + \frac{v_0^2}{c^2\alpha} + \alpha$:

(30)
$$\alpha B = \begin{pmatrix} b_1 & 0 & b_2 \\ 0 & \frac{1}{c^2}b_1 & \frac{1}{c^2}b_3 \\ b_2 & \frac{1}{c^2}b_3 & b_4 \end{pmatrix}.$$

The constraints (14) and (15) can be stacked in a $2n \times 4$-matrix V. It is

(31) $$h_1^t B h_2 = \left(h_{11}h_{12} + \frac{1}{c^2}h_{21}h_{22}, h_{12}h_{h31} + h_{11}h_{32}, \frac{1}{c^2}(h_{22}h_{31} + h_{32}h_{21}), h_{31}h_{32}\right)b,$$

(32) $$h_1^t B h_1 = \left(h_{11}^2 + \frac{1}{c^2}h_{21}^2, 2h_{11}h_{31}, \frac{2}{c^2}h_{21}h_{31}, h_{31}^2\right)b,$$

(33) $$h_2^t B h_2 = \left(h_{12}^2 + \frac{1}{c^2}h_{22}^2, 2h_{12}h_{32}, \frac{2}{c^2}h_{22}h_{32}, h_{32}^2\right)b.$$

But, instead of determining $\min \|Vb\|$ subject to $\|b\| = 1$ we are now able to formulate a meaningful constraint for the solution:

$$\frac{b_2^2}{b_1} + \frac{b_3^2}{c^2 b_1} + \frac{1}{b_1} = \frac{u_0^2}{\alpha} + \frac{v_0^2}{c^2\alpha} + \alpha = b_4 \tag{34}$$

must hold for a valid solution of the calibration problem. This means

$$b_2^2 + \frac{1}{c^2} b_3^2 + 1 = b_1 b_4 \Leftrightarrow b_1 b_4 - b_2^2 - \frac{1}{c^2} b_3^2 = 1 \tag{35}$$

should be fulfilled by a solution b^*. This can be put into matrix form by

$$b^t \underbrace{\begin{pmatrix} 0 & 0 & 0 & \frac{1}{2} \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -\frac{1}{c^2} & 0 \\ \frac{1}{2} & 0 & 0 & 0 \end{pmatrix}}_{=:C} b = 1. \tag{36}$$

The problem to determine $\min \|Vb\|^2 = \min b^t V^t V b$ subject to $b^t C b = 1$ can be solved by the Lagrange approach: A solution must satisfy

$$\left(V^t V - \lambda C\right) b = 0 \tag{37}$$

for a Lagrange-multiplier $\lambda \in \mathbb{R}$ (see e. g. [Jah96]). The task to find a suitable λ and b defines a generalized Eigenvalue problem. But, since C is regular, we achieve

$$\left(C^{-1} V^t V - \lambda I\right) b = 0. \tag{38}$$

Hence, the solution b^* is the Eigenvector associated to the smallest non-negative Eigenvalue of $C^{-1} V^t V$ (see e. g. [Bjö96]).

4.5 A starting solution with known aspect ratio and unknown skew

If we want to include the skewness γ into the projection matrix with a known aspect ratio c, we obtain

$$\tilde{P} = \begin{pmatrix} \alpha & \gamma & u_0 \\ 0 & c\alpha & v_0 \\ 0 & 0 & 1 \end{pmatrix} \tag{39}$$

and subsequently

$$B = \tilde{P}^{-t} \tilde{P}^{-1} = \begin{pmatrix} b_1 & b_4 & b_2 \\ b_4 & b_5 & b_6 \\ b_2 & b_6 & b_3 \end{pmatrix} \tag{40}$$

with

$$b_1 := \frac{1}{\alpha^2}, \qquad b_4 := -\frac{\gamma}{\alpha^3 c},$$
$$b_2 := \frac{v_0\gamma - u_0 c\alpha}{\alpha^3 c}, \qquad b_5 := \frac{\gamma^2}{\alpha^2 c^2} + \frac{1}{\alpha^2 c^2}, \tag{41}$$
$$b_3 := \frac{(u_0 c\alpha - v_0\gamma)^2}{\alpha^4 c^2} + \frac{v_0^2}{\alpha^2 c^2} + 1, \text{ and} \qquad b_6 := \frac{u_0 c\alpha\gamma - v_0\gamma^2}{\alpha^4 c^2} - \frac{v_0}{\alpha^2 c^2}.$$

A vector b^* minimizing $\|Vb\|$ with $b = (b_1, \ldots, b_6)^t$ and V derived from the constraints (14) and (15) as in the previous sections should satisfy

$$b_1 b_3 - b_2^2 = \frac{v_0^2}{\alpha^4 c^2} + \frac{1}{\alpha^2} > 0. \tag{42}$$

Since b^* defines parameters of the projection matrix $\tilde{P}$ only up to a scale factor, we demand

$$b_1 b_3 - b_2^2 = 1 \tag{43}$$

as a constraint for the vector minimizing $\|Vb\|$. This constraint can be expressed by a matrix $C \in \mathbb{R}^{6\times 6}$ with

$$C = \begin{pmatrix} C_1 & 0 \\ 0 & 0 \end{pmatrix} \text{ with } C_1 = \begin{pmatrix} 0 & 0 & \frac{1}{2} \\ 0 & -1 & 0 \\ \frac{1}{2} & 0 & 0 \end{pmatrix}. \tag{44}$$

We decompose $V^t V \in \mathbb{R}^{6\times 6}$ into three sub-matrices $S_1, S_2, S_3 \in \mathbb{R}^{3\times 3}$ such that

$$V^t V = \begin{pmatrix} S_1 & S_2 \\ S_2^t & S_3 \end{pmatrix} \tag{45}$$

holds. Then, the vector b^* minimizing $\|Vb\|$ subject to $b^t C b = 1$ is the Eigenvector to the smallest non-negative Eigenvalue of

$$M = C_1^{-1}(S_1 - S_2 S_3^{-1} S_2^t). \tag{46}$$

For a detailed derivation of the mathematics see [HF98].

4.6 A starting solution with no skew

For an unknown aspect ratio we must estimate both scaling factors α and β in the projection matrix $\tilde{P}$. Since γ should be small for actual camera mappings, we assume $\gamma = 0$ and therefore obtain a simpler problem than the one discussed in 4.1. With this the matrix $B = \tilde{P}^{-t}\tilde{P}^{-1}$ is given by

$$B = \begin{pmatrix} \frac{1}{\alpha^2} & 0 & -\frac{u_0}{\alpha^2} \\ 0 & \frac{1}{\beta^2} & -\frac{v_0}{\beta^2} \\ -\frac{u_0}{\alpha^2} & -\frac{v_0}{\beta^2} & \frac{u_0^2}{\alpha^2} + \frac{v_0^2}{\beta^2} + 1 \end{pmatrix}. \tag{47}$$

4.6.1 A straight forward constraint

Setting $b_1 := \frac{1}{\alpha^2}, b_2 := \frac{1}{\beta^2}, b_3 := -\frac{u_0}{\alpha^2}, b_4 := -\frac{v_0}{\beta^2}$, and $b_5 := \frac{u_0^2}{\alpha^2} + \frac{v_0^2}{\beta^2} + 1$ we parametrize

$$B = \begin{pmatrix} b_1 & 0 & b_3 \\ 0 & b_2 & b_4 \\ b_3 & b_4 & b_5 \end{pmatrix} \tag{48}$$

With the constraints (14) and (15) we obtain

$$\begin{aligned} h_1^t B h_2 &= \big(h_{11}h_{12}, h_{21}h_{22}, h_{12}h_{h31} + h_{11}h_{32}, h_{22}h_{31} + h_{32}h_{21}, h_{31}h_{32}\big)b, && (49) \\ h_1^t B h_1 &= \big(h_{11}^2, h_{21}^2, 2h_{11}h_{31}, 2h_{21}h_{31}, h_{31}^2\big)b, && (50) \\ h_2^t B h_2 &= \big(h_{12}^2, h_{22}^2, 2h_{12}h_{32}, 2h_{22}h_{32}, h_{32}^2\big)b. && (51) \end{aligned}$$

Since these equations are linear in b they can can be stacked in matrix form Vb with $V \in \mathbb{R}^{n\times 5}$.

An obvious constraint for the minimizing $b = (b_1, b_2, b_3, b_4, b_5)^t$ arises from B_{33}:

$$\frac{b_3^2}{b_1} + \frac{b_4^2}{b_2} + 1 = \frac{u_0^2}{\alpha^2} + \frac{v_0^2}{\beta^2} + 1 = b_5. \tag{52}$$

Unfortunately, this defines no quadratic constraint.

4.6.2 A least squares problem with quadratic constraint

To overcome the problem of a non-quadratic constraint, we could try to parametrize B by a vector of six values. Let $b = (b_1, \dots, b_6)^t \in \mathbb{R}^6$ be defined as $b_1 := \frac{1}{\alpha^2}, b_2 := \frac{1}{\beta^2}, b_3 := -\frac{u_0}{\alpha^2}, b_4 := -\frac{v_0}{\beta^2}, b_5 := \frac{u_0^2}{\alpha^2},$, and $b_6 = \frac{v_0^2}{\beta^2}$. Then we get

$$B = \begin{pmatrix} b_1 & 0 & b_3 \\ 0 & b_2 & b_4 \\ b_3 & b_4 & b_5 + b_6 + 1 \end{pmatrix}. \tag{53}$$

This parametrization introduces two obvious constraints for the entries of b:

$$\begin{aligned} b_5 &= \frac{b_3^2}{b_1} \Rightarrow b_5 b_1 = b_3^2, && (54) \\ b_6 &= \frac{b_4^2}{b_2} \Rightarrow b_6 b_2 = b_4^2. && (55) \end{aligned}$$

Therefore, $b_3^2 + b_4^2 - b_5 b_1 - b_6 b_2 = 0 \Leftrightarrow b^t C b = 0$ holds with

$$C = \begin{pmatrix} 0 & 0 & 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 & -\frac{1}{2} \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ -\frac{1}{2} & 0 & 0 & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{56}$$

Note that C has full rank.

With this parametrization the constraints (14) and (15) become

$$
\begin{aligned}
(57)\quad h_1^t B h_2 &= (h_{11}h_{12}, h_{22}h_{21}, h_{12}h_{31} + h_{11}h_{32}, h_{22}h_{31} + h_{21}h_{32}, h_{31}h_{32}, h_{31}h_{32})b \\
&\quad + h_{31}h_{32} = 0, \\
(58)\quad h_1^t B h_1 &= (h_{11}^2, h_{21}^2, 2h_{11}h_{31}, 2h_{21}h_{31}, h_{31}^2, h_{31}^2)b + h_{31}^2 = \\
h_2^t B h_2 &= (h_{12}^2, h_{22}^2, 2h_{12}h_{32}, 2h_{22}h_{32}, h_{32}^2, h_{32}^2)b + h_{32}^2.
\end{aligned}
$$

So, we obtain two equations which are linear in b. Stacking these equations determines a least squares problem $\|V'b - h'\|$ with $2n$ lines for n observations. The task is now to determine $b^* \in \mathbb{R}^6$ such that b^* minimizes $\|V'b - h'\|$ subject to $b^tCb = 0$. But, as one can see by the definition of V' the vectors $b' = (0,0,0,0,0,1)^t$ and $b'' = (0,0,0,0,1,0)^t$ are always solutions of the problem since it is $V'b' = V'b'' = h'$.

4.6.3 Combining necessary conditions to a quadratic constraint

We now return to the original parametrization of B by five parameters as in (48) with $b_1 := \frac{1}{\alpha^2}, b_2 := \frac{1}{\beta^2}, b_3 := -\frac{u_0}{\alpha^2}, b_4 := -\frac{v_0}{\beta^2}$, and $b_5 := \frac{u_0^2}{\alpha^2} + \frac{v_0^2}{\beta^2} + 1$. For a valid projection matrix the following inequalities must hold

$$(59)\qquad b_1b_2 > 0, b_1b_5 > 0, b_2b_5 > 0, \text{ and } b_3b_4 > 0$$

because these entries of B have the same sign. Hence,

$$(60)\qquad b_1b_2 + b_1b_5 + b_2b_5 + b_3b_4 > 0$$

must hold necessarily. Since the constraints (14) and (15) define a solution b only up to a scale factor, we demand that

$$(61)\qquad b_1b_2 + b_1b_5 + b_2b_5 + b_3b_4 = 1$$

should hold for a minimum of $\|Vb\|$ with V as defined by Equation 49 - 51. The quadratic constraint in Equation 61 can be put into a regular matrix $C \in \mathbb{R}^{5\times5}$ with

$$(62)\qquad C = \begin{pmatrix} 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \end{pmatrix}$$

by $b^tCb = 1$. Since C is regular the problem to minimize $\|Vb\|$ subject to $b^tCb = 1$ can be solved by finding the Eigenvector to the smallest positive Eigenvalue of $C^{-1}V^tV$ (see [RS06]).

The Eigenvalues of C are $-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}$, and 1. Since V^tV has only positive Eigenvalues, there are only two positive Eigenvalues of $C^{-1}V^tV$. This simple observation can be very helpful in a numerical aspect of the problem: An implementation may result in a situation with four or two negative Eigenvalues. In this case, the second highest Eigenvalue represents the solution.

4.6.4 A solution by a linear least squares problem with a Cholesky decomposition

If we set $\beta = c\alpha$ for an unknown factor $c \in \mathbb{R}_+$ we obtain

$$(63)\qquad \begin{pmatrix} \frac{1}{\alpha^2} & 0 & -\frac{u_0}{\alpha^2} \\ 0 & \frac{1}{c^2\alpha^2} & -\frac{v_0}{c^2\alpha^2} \\ -\frac{u_0}{\alpha^2} & -\frac{v_0}{c^2\alpha^2} & \frac{u_0^2}{\alpha^2} + \frac{v_0^2}{c^2\alpha^2} + 1 \end{pmatrix} = \frac{1}{\alpha^2 c^2} \begin{pmatrix} c^2 & 0 & -c^2 u_0 \\ 0 & 1 & -v_0 \\ -c^2 u_0 & -v_0 & c^2 u_0^2 + v_0^2 + c^2\alpha^2 \end{pmatrix}.$$

Since we are only able to determine a solution up to a scale factor, we set $b = (b_1, b_2, b_3, b_4)^t$ with $b_1 := c^2, b_2 := -c^2 u_0, b_3 := -v_0$, and $b_4 := c^2 u_0^2 + v_0^2 + c^2\alpha^2$ to parametrize B as

$$(64)\qquad \frac{1}{\alpha^2 c^2} \begin{pmatrix} b_1 & 0 & b_2 \\ 0 & 1 & b_3 \\ b_2 & b_3 & b_4 \end{pmatrix}$$

With the constraints (14) and (15) we obtain

$$(65)\qquad h_1^t B h_2 = (h_{11}h_{12}, h_{12}h_{31} + h_{11}h_{32}, h_{22}h_{31} + h_{21}h_{32}, h_{31}h_{32})b + h_{22}h_{21} = 0$$

$$(66)\qquad \begin{aligned} h_1^t B h_1 &= (h_{11}^2, 2h_{11}h_{31}, 2h_{21}h_{31}, h_{31}^2)b + h_{21}^2 = \\ h_2^t B h_2 &= (h_{12}^2, 2h_{12}h_{32}, 2h_{22}h_{32}, h_{32}^2)b + h_{22}^2. \end{aligned}$$

which can be stacked into a least squares problem $\|V'b - h'\|$ without additional constraints, since the entries of b are independent. A solution b^* of this linear leastsquares problem can be obtained by Householder transformations. If the matrix λB described by the solution b^* is positive definite, a Cholesky decomposition results in

$$(67)\qquad \lambda B = \lambda L^t L$$

Therefore, it is $L = \frac{1}{\lambda} P^{-1}$. The scale factor λ can be computed easily, since we know $(P^{-1})_{33} = P_{33} = 1$. If b^* does not describe a positive definite matrix we achieve only $u_0 = -\frac{b_2}{b_1}$ by the definition of the solution vector b^*, since b^* is only defined up to a scale factor λ. To obtain the remaining parameters α, β, and v_0 one has to determine λ.

With $\alpha = \frac{1}{c}\beta$ we derive

$$(68)\qquad \begin{pmatrix} \frac{c^2}{\beta^2} & 0 & -\frac{u_0 c^2}{\beta^2} \\ 0 & \frac{1}{\beta^2} & -\frac{v_0}{\beta^2} \\ -\frac{u_0 c^2}{\beta^2} & -\frac{v_0}{\beta^2} & \frac{u_0^2 c^2}{\beta^2} + \frac{v_0^2}{\beta^2} + 1 \end{pmatrix} = \frac{c^2}{\beta^2} \begin{pmatrix} 1 & 0 & -u_0 \\ 0 & \frac{1}{c^2} & -\frac{1}{c^2} v_0 \\ -u_0 & -\frac{1}{c^2} v_0 & u_0^2 + \frac{1}{c^2} v_0^2 + \frac{\beta^2}{c^2} \end{pmatrix}.$$

as a dual least squares problem. It allows us to determine v_0 instead of u_0 by a solution of this problem. v_0 can now be applied to determine the scale factor λ for b^* to obtain the remaining camera parameters.

5 Experimental results

5.1 Analyzed methods

In this section we show the performance of the proposed additional constraints on simulated and real data. We refer to the method of Zhang as presented in [Zha98] by **Zhang**. Also we denote **Zhang5** resp. **Zhang4** when we refer to the minimum with no skew ($\gamma = 0$) resp. no skew and identical scale factors in the projection matrix ($\gamma = 0.\alpha = \beta$). The first can be parametrized with four, the second with five parameters. Both parametrizations have been proposed by Zhang in [Zha98] to handle degenerate configurations.

Furthermore, we denote **OpenCV** for the closed form solution which is used in the Open Computer Vision Library (Open CV). The sources are available at sourceforge[1]. Unlike proposed in section 4.2 we applied no additional constraint. In contrast to the original code we apply Householder transformations instead of the pseudo-inverse to solve the linear least squares problem.

For the method with a known aspect ratio c and no skew as presented in section 4.4 we denote **Const4**. For the problem of finding a starting solution with no skew (section 4.6) we examine the solution with the necessary condition that some entries share the same sign (section 4.6.3) (denoted by **Const5**) and the least squares solution as introduced in section 4.6.4. We denoted this approach with **LSQ**. In our experiments we always applied the Cholesky decomposition to determine the projection matrix for this approach.

5.2 Simulations

To test the different starting solutions with valid reference data we use simulated data. Given a projection matrix and a calibration pattern at different positions we project the points by the camera mapping and simulate a pixel discretization with distortion. Thus, we are able to measure the difference of the obtain projection matrix to the reference mapping. In our tests we measured the Euclidean distance of the projection center to the reference center (referred to as "Center Error") and the Euclidean distance of the obtained scale factors to the reference factors (referred to as "Scale Error").

The minimization of $\|Vb\|$ with additional constraints defines an algebraic error. This means that the minimized error can not be seen in the image. Nonetheless, we apply the error of the standard camera calibration (7) to measure the quality of our proposed closed form solutions (referred to as "Error").

[1]Available at http://sourceforge.net/projects/opencvlibrary/

5.2.1 Simulation of the Siemens range camera

We show the gain of the additional constraints at simulated data derived from modelling the Siemens range camera as presented in [MLK+07]. In our simulation we model the image acquisition by a projection matrix with an aspect ratio of 130μm : 300μm. We assume an image with 64 pixels. Furthermore, we assume that we can extract a position of an observation with an accuracy of $\frac{1}{10}$ of a pixel from the pixel image. Each pixel position obtained by a projection of a model with 3×3 points is distorted by a Gaussian noise with mean 0 and variance σ. Tab. 1, 2, and 3 show the measured errors for different variances. We performed 1000 tests with the model at three different position (parallel, tilted at 0.2rad to the front, tilted at 0.2rad to the left) at a distance of 1m. Due to the noise not every experiment leads to a valid solution of the calibration problem. Hence, we counted these failures (referred to as "Misses").

It should be considered that we set the real principal point (u_0, v_0) for the method **OpenCV**, so that value for the entry "Center Error" must be always zero.

	Error	Center Error	Scale Error	Misses
Zhang	2.290444e+05	1.885655e+02	3.526991e+02	9
Zhang4	1.132229e+05	1.568470e+02	2.886693e+02	0
Zhang5	1.372039e+05	1.732284e+02	3.170649e+02	1
OpenCV	6.421307e+04	0	4.584806e+02	0
Const4	5.475627e+02	1.315052e+01	1.615813e+02	0
Const5	6.827966e+02	4.152875e+00	1.164377e+01	0
LSQ	3.663796e+02	7.947432e+00	9.654681e+01	0

Table 1: Pure pinhole camera with $\alpha = 120, \beta = 26, \gamma = 0, u_0 = 24, v_0 = 4$. 1000 tests with an image distortion with variance 0.5

	Error	Center Error	Scale Error	Misses
Zhang	4.078436e+04	4.103417e+01	1.209489e+02	25
Zhang4	1.818509e+05	7.904891e+01	1.366925e+02	13
Zhang5	4.067381e+04	4.365336e+01	1.224068e+02	25
OpenCV	3.356776e+04	0	1.685613e+02	0
Const4	5.261600e+04	2.587584e+01	1.712631e+02	0
Const5	2.351673e+03	4.089849e+00	7.675174e+01	0
LSQ	2.370006e+03	6.350904e+00	1.022842e+02	0

Table 2: Pure pinhole camera with $\alpha = 120, \beta = 26, \gamma = 0, u_0 = 24, v_0 = 4$. 1000 tests with an image distortion with variance 1

	Error	Center Error	Scale Error	Misses
Zhang	4.894989e+05	4.109475e+01	9.804577e+01	24
Zhang4	3.056477e+05	2.608063e+01	5.929147e+01	14
Zhang5	4.949557e+05	4.155502e+01	9.828731e+01	23
OpenCV	4.070450e+07	0	5.430675e+02	0
Const4	3.079750e+04	7.825065e+00	1.917129e+02	0
Const5	3.626391e+04	2.836785e+00	6.052411e+01	13
LSQ	2.619257e+04	6.647815e+00	9.979276e+01	0

Table 3: Pure pinhole camera with $\alpha = 120, \beta = 26, \gamma = 0, u_0 = 24, v_0 = 4$. 1000 tests with an image distortion with variance 1.5

One can see in Tab. 1, 2, and 3 that for increasing noise the proposed methods provide better results than the standard approach. In particular **OpenCV**, **Const4**, **Const5**, and **LSQ** deliver always valid solutions, but the results of **OpenCV** are much worse than **Const4**, **Const5**, and **LSQ** for images disturbed by a higher noise.

5.2.2 Simulation of a common camera

For a standard camera setup we set $\alpha = 700, \beta = 600, \gamma = 0, u_0 = 320, v_0 = 240$. We assume an image resolution of 640×480 pixels. Again, we assume an accuracy of point extraction of $\frac{1}{10}$ of a pixel. Each pixel position obtained by a projection of a model with 3×3 points is distorted by a Gaussian noise with mean 0 and variance σ. Table 4 and 5 show the measured errors for different variances. We performed 1000 tests with the model at the same three positions as for the test in section 5.2.1.

Most lenses in computer vision applications are assumed to be rotationally invariant. Therefore, it is obvious that aberrations caused by the lens depend only on the distance to the rotation axis (i. e. the optical axis). Hence, the radial distortion

(69)
$$\delta_r \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u + u \sum_{i=1}^{D} k_i (u^2 + v^2)^i \\ v + v \sum_{i=1}^{D} k_i (u^2 + v^2)^i \end{pmatrix}$$

with parameters $k_1, \ldots, k_D$ is commonly accepted to be dominating the observed distortion. Thus, we also applied a radial distortion in our simulated camera in Table 6.

The results in Tab. 4 and Tab. 5 confirm the ones of section 5.2.1: For higher distorted images the proposed methods deliver better results. This is also true for cameras with significant radial aberrations. The best performance of the approach **OpenCV** in Table

	Error	Center Error	Scale Error	Misses
Zhang	1.967200e+02	3.324624e+00	2.209671e+01	0
Zhang4	1.149798e+02	3.632442e+00	4.030154e+01	0
Zhang5	1.966540e+02	3.288516e+00	2.183596e+01	0
OpenCV	1.628845e+02	0	9.920800e+00	0
Const4	1.957935e+02	3.627571e+00	2.188091e+01	0
Const5	4.723808e+01	2.778486e+00	1.580470e+01	0
LSQ	3.238275e+02	3.275714e+00	2.171136e+01	0

Table 4: Pure pinhole camera with $\alpha = 700, \beta = 600, \gamma = 0, u_0 = 320, v_0 = 240$. 1000 tests with an image distortion with variance 0

	Error	Center Error	Scale Error	Misses
Zhang	1.247490e+05	4.085757e+01	1.998144e+02	0
Zhang4	9.796890e+04	1.649974e+01	1.129240e+02	0
Zhang5	9.861288e+04	1.646876e+01	9.712332e+01	0
OpenCV	1.040106e+05	0	4.549401e+01	0
Const4	9.720122e+04	8.080718e+00	4.554056e+01	0
Const5	9.829203e+04	4.765583e+00	2.075606e+01	0
LSQ	9.869905e+04	9.151808e+00	4.792062e+01	0

Table 5: Pure pinhole camera with $\alpha = 700, \beta = 600, \gamma = 0, u_0 = 320, v_0 = 240$. 1000 tests with an image distortion with variance 1.0

6 must be seen in connection with the fact that the knowledge of the correct principal point might not be given for real camera setups.

5.3 Real data

5.3.1 A web camera

To calibrate a common web camera we used a calibration model with a regular grid of 20×19 points with a distance of 1cm to each other at four different positions. The resolution of the camera was 1028×768 pixels, which is not the physical but an interpolated resolution. Fig. 2 shows an exemplary input image of the data set.

Tab. 7 shows that all proposed methods deliver better results than the standard approach. But, the best result was obtained by the modified standard approach **Zhang5** which uses five instead of six parameters. For the method **Const4** we assumed $\alpha = \beta$, i.e. $c = 1$ in Equation 36.

	Error	Center Error	Scale Error	Misses
Zhang	1.011916e+06	1.563035e+02	1.114902e+02	0
Zhang4	6.445274e+05	1.052913e+02	8.659299e+01	0
Zhang5	1.847035e+06	2.677292e+02	4.066258e+02	0
OpenCV	2.497499e+05	0	1.699679e+01	0
Const4	6.341745e+05	9.639704e+01	2.263665e+01	0
Const5	8.043419e+05	1.271122e+02	9.344606e+01	0
LSQ	4.328916e+05	4.834713e+01	1.643554e+02	0

Table 6: Pinhole camera with $\alpha = 700, \beta = 600, \gamma = 0, u_0 = 320, v_0 = 240$, and radial distortion ($D = 2, k_1 = 1.0, k_2 = 0.3$) see Equation 69. 1000 tests with an image distortion with variance 1.0

	α	β	γ	u_0	v_0	Error
Zhang	1413.95491	1417.66028	1.38618	511.13524	395.05554	1.15213e+07
Zhang4	1405.81051	1405.81051	0	506.10194	396.65926	8.00958e+06
Zhang5	1412.56213	1412.56213	0	511.64761	395.58454	5.61085e+06
OpenCV	1411.03620	1415.18209	0	511.50000	383.50000	7.16744e+06
Const4	1405.68844	1405.68844	0	506.05717	396.61410	6.92353e+06
Const5	1414.24322	1418.10229	0	512.24748	395.78116	7.79540e+06
LSQ	1412.46609	1416.04116	0	511.60233	395.56954	1.10729e+07

Table 7: Result of the “webcam” data set

5.3.2 Wide angle lens

A wide angle lens introduces a different kind of distortion, the so called *fish-eye distortion*. Obviously, such a camera is far from being a pure pinhole camera (see Fig. 3). The assumptions made to obtain a closed form solution of the camera calibration problem are no longer valid. Nevertheless, we may apply the proposed methods to obtain a starting solution for a subsequent non-linear optimization, while the classic approach fails sometimes. We used two disjunct data sets of four images in our experiments observing the same model as in the previous section. Tab. 8 shows the result for the first data set: For this setup **Zhang** and **Zhang5** deliver no valid solution. For the second data set (Tab. 9) **Zhang4** fails. Moreover, the error values for the proposed approaches are in most cases better than the ones for the standard methods.

5.3.3 Common CCD camera with 8mm lens

For the last test we used a common camera with an $\frac{1}{3}''$ CCD sensor and an $8mm$ lens. Fig. 4 shows that this setup delivers images with low visible distortion. Again we used

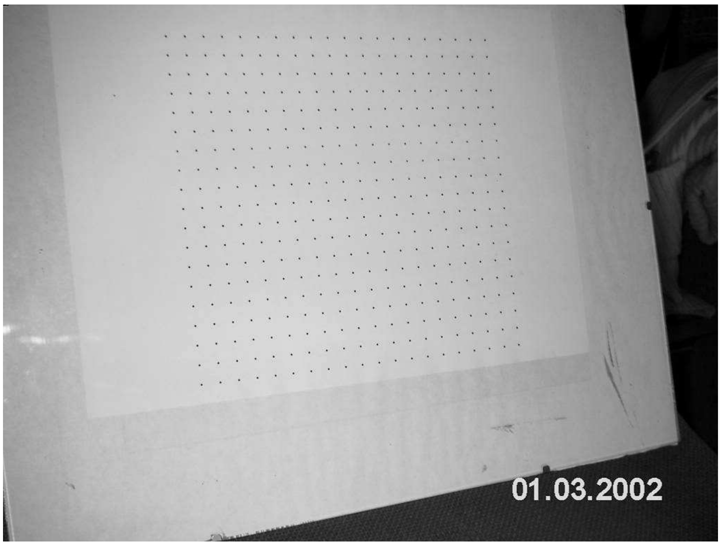

Figure 2: Input image for the “webcam” data set

	α	β	γ	u_0	v_0	Error
Zhang	No valid solution					
Zhang4	492.43687	492.43687	0	348.56854	1025.33368	4.19843e+07
Zhang5	No valid solution					
OpenCV	425.99208	442.95108	0	191.50000	143.50000	1.37736e+07
Const4	-382.42165	-382.42165	0	222.15850	111.63821	1.46295e+07
Const5	154.78194	155.23324	0	239.08388	249.32366	1.31352e+07
LSQ	295.43874	209.74102	0	222.72716	129.81869	1.55342e+07

Table 8: Result of the wide angle data set 1

a model with 20×19 points on an regular grid with a distance of 1cm to each other. In constrats to the two previous experiments we used ten positions of the model.

In this case the assumption that the principal point is the center of the image leads to the second best result (see Tab. 10). If we set $\alpha = \beta$, i.e. $c = 1$ for the constraint matrix C in (36) for the method **Const4** we obtain the best result. Note that this assumption might only approximately be true. Moreover, the proposed solutions **Const5** and **LSQ** deliver better results than the standard solution **Zhang**. Surprisingly, the approach **Zhang5** is better than **Const5** and **LSQ**.

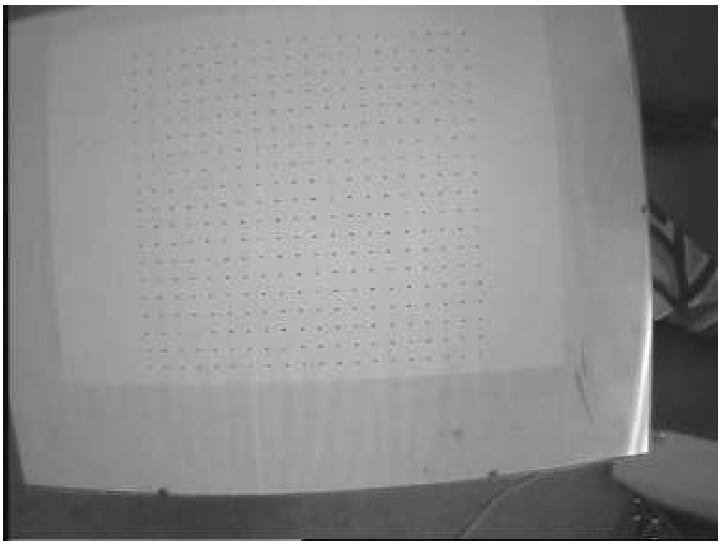

Figure 3: Input image for the “wide angle” data set

	α	β	γ	u_0	v_0	Error
Zhang	946.39841	1172.14759	-27.60928	-921.0860	246.91108	1.42441e+07
Zhang4	No valid solution					
Zhang5	955.44289	1169.00472	0	-863.20976	241.90690	1.71064e+07
OpenCV	650.38393	729.75649	0	191.50000	143.50000	1.74287e+06
Const4	-907.74050	-907.74050	0	514.83132	213.01392	5.19799e+06
Const5	580.77123	632.91195	0	-70.95787	204.37658	1.86308e+06
LSQ	168.52903	192.51794	0	115.39772	197.94121	1.35655e+06

Table 9: Result of the wide angle data set 2

6 Conclusion

In this article we investigated some additional constraints for the closed solution of the camera calibration problem. We were able to formulate additional constraints under different restrictions of the camera mapping. Some of these restrictions (like a known principal point) are quite common or - as in the case of a known aspect ratio - are easy to obtain for special hardware. But, even if the aspect ratio is unknown we are able to formulate a necessary quadratic side condition for a valid solution of the camera calibration problem.

The experimental results have shown that for degenerated configurations (i.e. wide angle lenses or image devices with low resolutions) the proposed additional constraints provide a valid solution of the closed form camera calibration problem, which is not

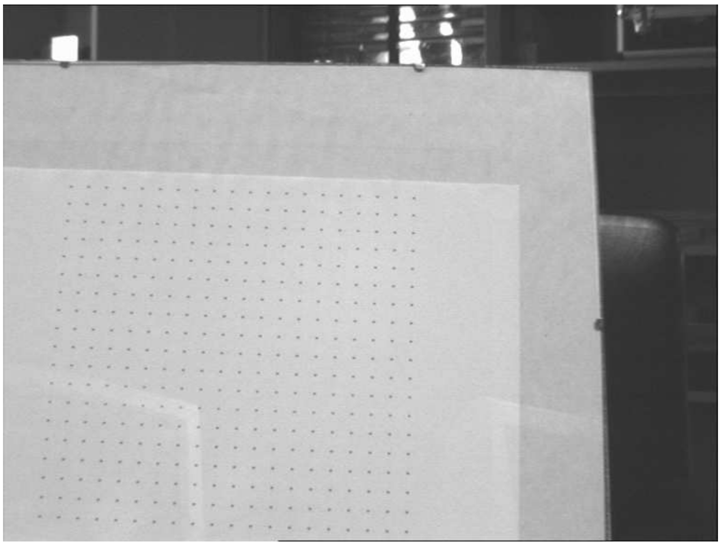

Figure 4: Input image for the “common” data set

	α	β	γ	u_0	v_0	Error
Zhang	1364.94803	1372.94762	-0.65236	340.20294	297.27872	2.06340e+07
Zhang4	1462.94617	1462.94617	0	353.23331	260.73553	2.00844e+07
Zhang5	1367.72411	1375.75560	0	341.04048	296.80323	1.90253e+07
OpenCV	1461.31676	1467.69818	0	383.50000	287.50000	1.65585e+07
Const4	1438.65592	1438.65592	0	349.90556	262.30899	5.19799e+06
Const5	1349.58029	1357.85751	0	339.04232	299.54251	1.95691e+07
LSQ	1335.25135	1343.38407	0	337.46678	301.68071	1.93194e+07

Table 10: Result of the “common” data set

true for the standard approach. Thus, such methods can be applied to obtain a starting solution for a subsequent non-linear optimization.

References

[Atk96] K.B. Atkinson, editor. *Close Range Photogrammetry and Machine Vision*. Whittle Publishing, 1996.

[Bjö96] Å. Björck. *Numerical Methods for Least Squares Problems*. SIAM, Philadelphia, 1996.

[Bra02] G.R. Bradski. OpenCV: Examples of use and new applications in stereo, recognition and tracking. In *Vision Interface*, page 347, 2002.

[Har97] R. Hartley. In defense of the eight-point algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(6):580–593, Juni 1997.

[HF98] R. Halir and J. Flusser. Numerically stable direct least squares fitting of ellipses. In V. Skala, editor, *Proc. Int. Conf. in Central Europe on Computer Graphics, Visualization and Interactive Digital Media*, pages 125 – 132, 1998.

[HS97] J. Heikkilä and O. Silven. A four-step camera calibration procedure with implicit image correction. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1106 – 1112, San Juan, Puerto Rico, 1997. IEEE.

[HSG06] T. Hanning, R. Schöne, and S. Graf. A closed form solution for monocular re-projective 3d pose estimation of regular planar patterns. In *International Conference of Image Processing (ICIP)*, pages 2197 – 2200, Atlanta, Georgia, 2006.

[HZ00] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.

[Jah96] Johannes Jahn. *Introduction to the Theory of Nonlinear Optimization*. Springer, 2 edition, 1996.

[MLK+07] P. Mengel, L. Listl, B. Koenig, M. Pellkofer, U. Wagner, and R. Wertheimer. Time-of-flight camera for pedestrian protection and collision mitigation. In *6th European Congress and Exhibition on Intelligent Transport Systems and Services*, 2007. accepted.

[RS06] A. Zimmermann R. Schöne, T. Hanning. Approximation mit quadratischer Nebenbedingung. Technical Report MIP-0605, Fakultät für Mathematik und Informatik, Universität Passau, 2006.

[Tsa87] R. Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Journal of Robotics and Automation*, 3(4):323 – 344, 1987.

[WCH92] J. Weng, P. Cohen, and M. Herniou. Camera calibration with distortion models and accuracy evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(10):965 – 980, 1992.

[WM94] G.-Q. Wei and S. De Ma. Implicit and explicit camera calibration: Theory and experiments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16:469 – 480, 1994.

[Zha98] Z. Zhang. A Flexible new technique for camera calibration. Technical report, Microsoft Research, 1998. Technical Report MSR-TR-98-71.