University of Passau
Technical Report

Euclidean vs. projective camera calibration: Algorithms and Effects on 3D-reconstruction

Tobias Hanning and Simone Graf *

September 13, 2007

Abstract

The camera mapping can be seen in two ways. The classic approach is to emphasize the projective nature of the camera. But also the re-projective nature can be taken into account: Every point in the image plane determines a viewing ray. Both mappings can be described by the same set of parameters. In fact the re-projective camera mapping can be seen as the inversion of the projective camera mapping. A calibration algorithm determines the parameters which describe the camera mapping in a non-linear optimization algorithm. In this article we compare two error functions: The projective error function measures the distance of the projected prototype to the observed points in the image plane. The re-projective error function measures the distance of the prototype to the re-projected rays, which are determined by the observed points. We present calibration algorithms considering distortions for both error functions and compare them with regard to the 3D-reconstruction problem.

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

Contents

1 Introduction

Most technical applications of computer vision need calibrated cameras. In particular 3D-reconstruction tasks depend crucially on the knowledge of the projection properties of the involved cameras. In this article we compare two approaches to calibrate cameras: On the one hand we examine the well known projective approach to camera calibration (cf. [Tsa87], [Zha98], [HS97], [WCH92]), on the other hand we formulate a re-projective approach.

Both methods define a nonlinear optimization problem to determine the optimal camera. So the first step is the parametrization of the set of cameras to formulate the optimization problem. The next step towards an optimization routine is to establish an error function for the given problem. On this topic the two approaches differ: The projective camera calibration measures the distance of the observed points in the image plane to the projected points of the calibration pattern. This means that the distances are measured in the image coordinate system. So the unit of the distance is “pixel”.

In [HGP04] we showed that in the area of stereo camera vision the re-projective approach to the camera mapping allows a simpler handling of the so called epipolar constraint. We used a classic projective algorithm to calibrate the cameras and derived the re-projective properties by inversion of the camera mapping. This approach encourages to take a closer look at the re-projective nature of cameras. Since in [HGP04] we only used the re-projective mapping, we should be able to calibrate our camera w.r.t. an error function derived from the re-projection problem. This re-projective calibration takes advantage of the “image point to viewing ray” property of the camera mapping: All points that will be projected on the same point in the image plane determine a straight line in space. Conversely each point in the image plane determines a straight line in the reference coordinate system of $\mathbb{R}^3$. The re-projective calibration measures the Euclidean distance of calibration points to the re-projected lines of their observations. Since this measurement is done in the reference coordinate system the unit of this distances is metric in the case of a metric Cartesian reference coordinate system.

In the next section we define our camera model and its parametrization. This parametrization allows us to formulate the error functions for both approaches in the third section. In section four we present two methods of 3D-reconstruction. To have equal conditions when comparing the two approaches, we must present two ways of Euclidean reconstruction either: First, a method minimizing the projective error function to ascertain the 3D-point, which we call *projective reconstruction*, second a method taking the re-projective error function into account, which we call *re-projective reconstruction* or *Euclidean reconstruction*. This finally allows us to compare the two calibration methods in the fifth section. A discussion and conclusion will end this article.

2 Modeling the camera

2.1 The camera mapping

In the following the expression "camera" denotes the whole image acquisition system including the camera, the lens and its position in a reference coordinate system.

Let $\mathfrak{K}$ be the set of all cameras. Each element $\mathcal{K} \in \mathfrak{K}$ defines a mapping $\mathcal{K} : \mathbb{R}^3 \to \mathbb{R}^2$. We call this mapping the *projection*. To solve Euclidean reconstruction problems we are also interested in the "inverse" mapping $\mathcal{K}^{-1} : \mathbb{R}^2 \to \mathbb{R}^3$. Obviously for $i \in \mathbb{R}^2$ the set $\vec{i}_{\mathcal{K}} := \{x \in \mathbb{R}^3 | \mathcal{K}(x) = i\}$ is in general not a singleton. Like most approaches to camera calibration we model our camera mapping as a pinhole camera with distortion. Since we model our camera w.r.t. geometric optics we can assume that $\vec{i}_{\mathcal{K}}$ is a straight line ([Hec87]). We call the mapping $i \mapsto \vec{i}_{\mathcal{K}}$ the *re-projection* of the camera mapping. For a simpler notation we identify $\vec{i}_{\mathcal{K}}$ with $\mathcal{K}^{-1}(\{i\})$. For the pinhole assumption all lines $\mathcal{K}^{-1}(\{i\})$ (for $i \in \mathbb{R}^2$) intersect in one point called the *pinhole*.

2.2 Camera parametrization

We subsequently present two approaches to parameterize the set of all cameras $\mathfrak{K}$ of the pinhole model with distortions. We distinguish two approaches to describe the camera mapping: First a complete parametrization in Euclidean coordinates is given, second we present the well known approach to describe the camera mapping as a projective mapping. This has the advantage that the main part of the mapping can be encoded by matrices.

2.2.1 Camera parametrization in Euclidean coordinates

The first step of every camera mapping is a coordinate transformation from the reference coordinate system to the camera coordinate system. All parameters describing this coordinate transformation $T : \mathbb{R}^3 \to \mathbb{R}^3$ with $p \mapsto Rp + t$ where $R \in \mathbb{R}^{3\times 3}$ is a rotation matrix and $t \in \mathbb{R}^3$ a translation vector are called *extrinsic parameters*. Since the rotation can be defined by three angles the extrinsic parameters consists of six real values (Euler's rotation theorem).

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

$$\mathrm{P}_z : \begin{array}{c} \mathbb{R}^3 \setminus \{z = 0\} \to \mathbb{R}^2 \\ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \mapsto \begin{pmatrix} x/z \\ y/z \end{pmatrix} \end{array} .$$

Note that P_z describes no change of units. The result of P_z still describes a point in the camera coordinate system.

The distortion mapping $\delta : \mathbb{R}^2 \to \mathbb{R}^2$ is defined in the image plane with respect to the camera coordinate system. The most common distortion model is the one of a radial distortion

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

with parameters $k_1, \ldots, k_D$ where in most cases is $D = 2$.

The last step accomplishes the change of the camera coordinate system to the (computer) image coordinate system. This is also a change of units: metric to pixel. We set

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

where (u_0, v_0) is the optical center with respect to the pixel coordinate system. If f is the focal length of the camera and $d_u \times d_v$ is the dimension of a CCD-element, then the parameters α and β can be interpreted as $\alpha = \frac{f}{d_u}$ and $\beta = \frac{f}{d_v}$. Some authors (e.g. [Tsa87]) prefer this more physical exemplification of the matrix P. γ describes the skewness between the axes of the pixel coordinate system. If γ is zero the coordinate axes of the image coordinate system are perpendicular.

After all, our camera mapping $\mathcal{K}$ can be parameterized by $\mathcal{K} = P \circ \delta \circ \mathrm{P}_z \circ T$. All parameters which describe the mapping $\Pi := P \circ \delta \circ \mathrm{P}_z$ are called *intrinsic camera parameters*.

2.2.2 Camera parametrization in projective coordinates

The transition from a Euclidean reference coordinate system to a projective coordinate system allows us to denote changes of coordinate systems as matrices. Since the camera mapping consists of coordinate system transformations except the distortion, nearly the whole camera mapping can be described by a matrix.

We define

$$\tilde{T} = \begin{pmatrix} Rt \end{pmatrix} \in \mathbb{R}^{3\times 4} .$$

Using the notation of Zhang (see [Zha98]) we set

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

and $\tilde{\delta}((u, v, 1)^t) = (\delta(u, v), 1)^t$. Assuming an ideal camera without distortion we define

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

to describe the camera mapping. The projection $(u,v)^t$ of a world point $(x,y,z)^t$ by the camera mapping fulfills

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

for a scalar $s \in \mathbb{R}$. Equation 1 is called *pinhole model equation*. So for every world point $X = (x,y,z)^t$ the projection U of X by the camera $\mathcal{K}$ can be obtained by setting $\widetilde{U} = (\tilde{u}, \tilde{v}, \tilde{w}) = \tilde{\mathcal{K}}(x,y,z,1)^t$ and $U = (\frac{\tilde{u}}{\tilde{w}}, \frac{\tilde{v}}{\tilde{w}})^t$, if $\tilde{w} \neq 0$. Otherwise the world point has no image point and accordingly is a point at infinity in the projective sense.

3 Camera calibration

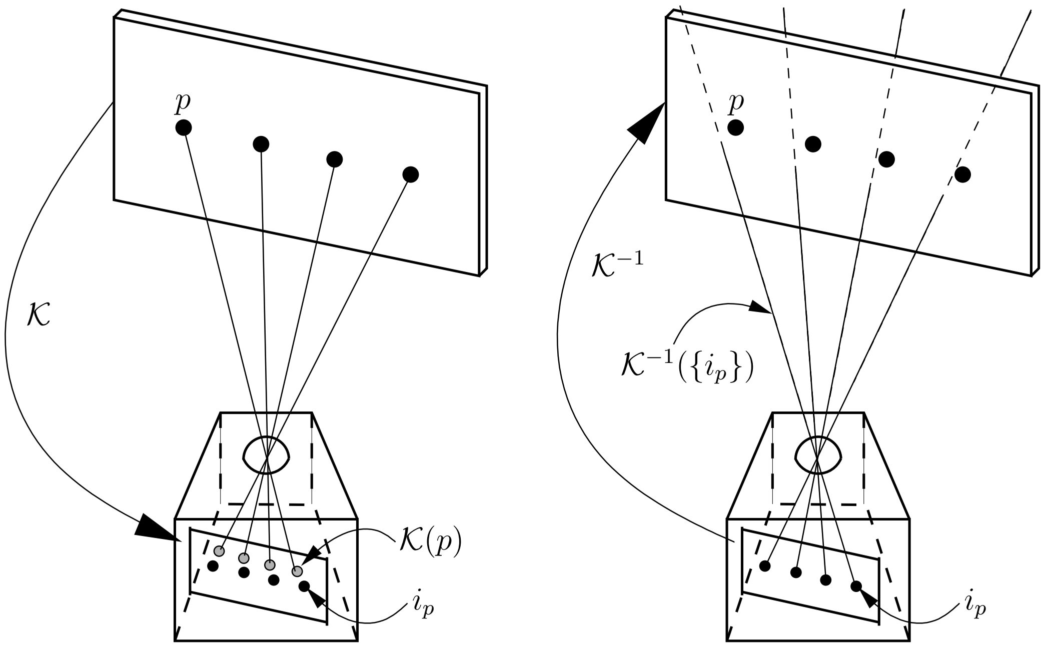
Let $\mathbf{P} \subset \mathbb{R}^3$ be a finite set of points with respect to the reference coordinate system. For every $p \in \mathbf{P}$ we denote $i_p \in I$ for the observed projection of p in the image plane I with respect to the image coordinate system.

3.1 Projective calibration

The *projective calibration* minimizes the function

$$\Phi : \begin{array}{l} \mathfrak{K} \to \mathbb{R}_+ \\ \mathcal{K} \mapsto \sum_{p \in \mathbf{P}} \|i_p - \mathcal{K}(p)\|^2 \end{array} . \tag{2}$$

This means that the projective calibration minimizes the distances of the observed image



(a) Projective calibration (b) Re-projective calibration

Figure 1: Minimized distances of the calibration methods

points i_p to the world points projected on the image plane by the camera mapping $\mathcal{K}$ (see Fig. 1(a)).

3.2 Re-projective calibration

As mentioned before each point in the image plane determines a straight line in the reference coordinate system intersecting the pinhole of the camera. The re-projected ray of a point $i \in I$ is defined by $\mathcal{K}^{-1}(\{i\})$.

Unfortunately $\mathcal{K}^{-1} = (P \circ \delta \circ \mathrm{P}_z \circ T)^{-1} = T^{-1} \circ \mathrm{P}_z^{-1} \circ \delta^{-1} \circ P^{-1}$ is not well defined: For $(u,v) \in \mathbb{R}^2$ the set $\mathrm{P}_z^{-1}(\{(u,v)^t\}) = \{s(u,v,1)^t | s \in \mathbb{R}\}$ is a straight line in $\mathbb{R}^3$ with direction $(u,v,1)^t$ containing the origin. In order to construct a well defined function P_z^{-1} we choose a suitable representative of $\mathrm{P}_z^{-1}(\{i\})$ by setting $\mathrm{P}_z^{-1}((u,v)^t) := \frac{1}{\sqrt{(u^2+v^2+1)}}(u,v,1)^t$. Note that this representative has norm 1. When we refer to this special representative of the pre-image of Π we denote $\Pi^{-1}(i)$ instead of $\Pi(\{i\})$.

Another difficulty in inverting the camera mapping $\mathcal{K}$ arises from the distortion function δ. For the radial distortion δ_r no analytical way to determine δ_r^{-1} exists. So one has to approximate δ_r^{-1} e.g. by an iteration process (see [PWH97] or [HS97]). To determine different parameters for the transformation from undistorted to distorted coordinates and from distorted to undistorted coordinates (see [TYO02] or [WM94]) provides another possibility to represent the inverse mapping. Methods to remove the distortion without estimating the specific parameters in images containing lines (see e.g. [DF95]) can also be found in literature.

The *re-projective calibration* minimizes the function

$$\Psi : \begin{array}{l} \mathfrak{K} \to \mathbb{R}_+ \\ \mathcal{K} \mapsto \sum\limits_{p \in \mathbf{P}} \mathrm{dist}_3(p, \mathcal{K}^{-1}(\{i_p\}))^2 \end{array} . \tag{3}$$

In this case "dist$_3$" denotes the distance of the world point to the re-projected ray which is determined by the observed point i_p (see Fig. 1(b)). It is the Euclidean distance of a point to a line.

With $\mathcal{K} = \Pi \circ T$ it is $\mathcal{K}^{-1} = T^{-1} \circ \Pi^{-1}$. Further for every point $p, q \in \mathbb{R}^3$, and every Rotation $R \in \mathbb{R}^{3\times 3}$, and translation $t \in \mathbb{R}^3$ it is $\|p - (Rq + t)\| = \|R^{-1}p - R^{-1}t - q\|$ thus we get $\|p - T(q)\| = \|T^{-1}(p) - q\|$.

For a straight line l we achieve $\mathrm{dist}_3(p, T^{-1}(l)) = \mathrm{dist}_3(T(p), l)$, which allows us to parameterize the line l only by its direction if l contains the origin.

It is $\mathrm{dist}_3(T(p), l)^2 = \min_{q \in l} \|T(p) - q\|^2 = \|T(p) - \hat{p}\|^2$ where $\hat{p}$ is the orthogonal projection of $T(p)$ on l. If l is a line through the origin then l can be parameterized by its direction $\vec{d}$ with $\|\vec{d}\| = 1$. If $\bullet$ denotes the inner product it is $\hat{p} = \left(\vec{d} \bullet T(p)\right)\vec{d}$. This leads us to a very effective method to determine $\mathrm{dist}_3(T(p), l)^2$ by

$$\mathrm{dist}_3(T(p), l)^2 = \|T(p)\|^2 - \left(\vec{d} \bullet T(p)\right)^2 .$$

For an image point $i_p \in I$ we get the corresponding direction by setting $\vec{d_{i_p}} = \Pi^{-1}(i_p)$. Hence the error function (3) of the re-projective calibration becomes

$$(T, \Pi) \mapsto \sum_{p \in \mathbf{P}} \|T(p)\|^2 - \left(\Pi^{-1}(i_p) \bullet T(p)\right)^2 . \tag{4}$$

3.3 Calibration with multiple targets

For flexible calibration purposes the set of points $\mathbf{P}$ is often a grid of points on a plane (see Fig. 2). Since the reference coordinate system is defined by the calibration plate the z-coordinate of each calibration points is zero. It is easy to see that both calibration methods yield no sufficient result when using only one image of such a calibration plate (see e.g. [Zha98]). On the other hand a planar calibration plate provides a great flexibility when using several images of a plate at different positions and angles. With multiple targets the number of parameters in our error functions (Eq. 2 and Eq. 3) rises: we have to determine the position of every plate. Let $\prod$ be the set of all cameras whose extrinsic parameters define the identity and $\mathcal{T}$ the set of all transformations. Then for N positions of the calibration plate the error functions become

$$\Phi_K : \begin{array}{l} \prod \times \mathcal{T}^N \to \mathbb{R}_+ \\ (\Pi, T_1, \ldots, T_N) \mapsto \sum_{c=1}^{N} \sum_{p \in \mathbf{P}} \| i_{c,p} - \Pi \circ T_c(p) \|^2 \end{array} \tag{5}$$

in the projective case and

$$\Psi_K : \begin{array}{l} \prod \times \mathcal{T}^N \to \mathbb{R}_+ \\ (\Pi, T_1, \ldots, T_N) \mapsto \sum_{c=1}^{N} \sum_{p \in \mathbf{P}} \mathrm{dist}_3(p, \Pi^{-1}(T_c\{i_{c,p}\}))^2 \end{array} \tag{6}$$

in the re-projective case. $i_{c,p}$ denotes the observation of the point $p \in \mathbf{P}$ in the c-th image of the calibration plate.

3.4 Implementation

Like all non-linear optimization algorithms we need a suitable starting value to achieve an acceptable solution of our camera calibration algorithms. Several attempts to obtain a closed form estimation from observed calibration patterns can be found in the literature. They differ from using the camera's manual to utilize constraints of projective mappings. We use the approach of Zhang [Zha98] to get the starting value of our non-linear optimization algorithms. Other suitable techniques can be found in [Tsa87], [WCH92] or [WM94].

We use the same optimization algorithms for the implementation of both approaches. They only differ in the error function of the non-linear optimization. As mentioned above we use the closed form solution of Zhang [Zha98] to obtain a starting value for our nonlinear optimization. The algorithm of Zhang works with multiple targets, so we have to calculate the transformation part (rotation and translation of the camera with respect to the reference coordinate system) for every target. The second step in our calibration algorithm is a refinement of these estimated transformations. Given the inner parameter estimated by the closed form solution we recalculate each position of our calibration targets using the Levenberg-Marquardt-algorithm [Mar63].

The third step is to refine the inner camera parameter with fixed transformations, and the fourth step is to refine all camera parameter by the non-linear optimization algorithm of Levenberg and Marquardt.

Until now we have not yet considered the distortion mapping δ in the calibration algorithm. For our implementation we choose Δ as the set of radial distortions with $D = 2$ and determine δ_r for fixed P and T by minimizing

$$\begin{aligned} &\Delta \to \mathbb{R}_+ \\ &\delta_r \mapsto \sum_{p \in \mathbf{P}} \|i_p - \underbrace{(P \circ \delta_r \circ \mathrm{P}_z \circ T)}_{\mathcal{K}}(p)\|^2 \,, \end{aligned} \tag{7}$$

which can be approximated by minimizing

$$\begin{aligned} &\Delta \to \mathbb{R}_+ \\ &\delta_r \mapsto \sum_{p \in \mathbf{P}} \|P^{-1}(i_p) - \delta_r\left(\left(\mathrm{P}_z \circ T\right)(p)\right)\|^2 \,. \end{aligned} \tag{8}$$

On the other hand we approximate δ_r^{-1} by minimizing

$$\begin{aligned} &\Delta \to \mathbb{R}_+ \\ &\delta_r' \to \sum_{p \in \mathbf{P}} \|\delta_r'\left(P^{-1}(i_p)\right) - \left(\mathrm{P}_z \circ T\right)(p)\|^2 \,. \end{aligned} \tag{9}$$

Note that in our implementation not only $\delta_r \in \Delta$ but also the inverse mapping δ_r' is in Δ. In general minimizing (9) does not result in δ_r^{-1}, but in a good approximation which is sufficient for practical use. As a matter of fact δ_r itself is only an approximation of the real distortion function.

The last step of our calibration algorithms is a non-linear optimization of all parameters by the Levenberg-Marquardt-algorithm with error function (2) for the projective calibration and (4) for the re-projective calibration.

4 3D-reconstruction

We compare the approaches to camera calibration by 3D-reconstruction problems. When we use a setup of two cameras observing the same scene a point in the reference coordinate system will be reconstructed by the observed projections of this point in images of both cameras.

Both approaches to camera calibration determine their own approach to this (stereo-) reconstruction problem. The projective reconstruction minimizes the distances of the projections of the calculated point to the observed points in the image planes. The re-projective reconstruction determines the point of least Euclidean distance to the re-projected rays to estimate the position of real point.

4.1 Projective reconstruction

The projective reconstruction is frequently accompanied by making use of the camera parametrization in projective coordinates as a matrix. To do so it is necessary to uncouple the distortion mapping: The distortion of the measured pixel coordinates must be removed using one of the techniques cited above.

It should be noted that this procedure models a different camera mapping $\mathcal{K}' = \delta_r' \circ P \circ \mathrm{P}_z \circ T$, because the inverse mapping first removes the distortion. Anyway this approach is commonly used, because replacing δ_r' with $P \circ \delta_r \circ P^{-1}$ maintains the original camera mapping, and the resulting undistorted coordinates permit us to regard the remaining camera mapping as ideal camera without distortion.

Reconstructing points can be achieved by a linear and a nonlinear method to estimate the best possible point of intersection by its projections. Let the 3D-coordinate in space be $X = (x, y, z)^t$ and its corresponding (undistorted) image coordinates be $m_1 = (u_1, v_1, 1)^t$ and $m_2 = (u_2, v_2, 1)^t$. Furthermore let $\tilde{\mathcal{K}}^{(1)}$ and $\tilde{\mathcal{K}}^{(2)}$ be the two cameras of a stereo setup. Then the following two equations can be defined by making use of the pinhole model equation (1) and the camera projection matrices related to the two images:

$$s_1(u_1, v_1, 1)^t = \tilde{\mathcal{K}}^{(1)}(x, y, z, 1)^t \quad (10)$$

$$s_2(u_2, v_2, 1)^t = \tilde{\mathcal{K}}^{(2)}(x, y, z, 1)^t \quad (11)$$

where s_1 and s_2 are two arbitrary scalars. They can be determined by the inner product of the third rows of the projection matrices:

$$\begin{aligned} s_1 &= (\tilde{\mathcal{K}}^{(1)}_{3,1}, \tilde{\mathcal{K}}^{(1)}_{3,2}, \tilde{\mathcal{K}}^{(1)}_{3,3})X + \tilde{\mathcal{K}}^{(1)}_{3,4} \\ \text{and } s_2 &= (\tilde{\mathcal{K}}^{(2)}_{3,1}, \tilde{\mathcal{K}}^{(2)}_{3,2}, \tilde{\mathcal{K}}^{(2)}_{3,3})X + \tilde{\mathcal{K}}^{(2)}_{3,4} \ . \end{aligned}$$

Furthermore the equations (10) and (11) can then be rewritten as

$$AX = d \ , \quad (12)$$

with

$$A = \begin{pmatrix} \tilde{\mathcal{K}}^{(1)}_{1,1} - u_1\tilde{\mathcal{K}}^{(1)}_{3,1} & \tilde{\mathcal{K}}^{(1)}_{1,2} - u_1\tilde{\mathcal{K}}^{(1)}_{3,2} & \tilde{\mathcal{K}}^{(1)}_{1,3} - u_1\tilde{\mathcal{K}}^{(1)}_{3,3} \\ \tilde{\mathcal{K}}^{(1)}_{2,1} - v_1\tilde{\mathcal{K}}^{(1)}_{3,1} & \tilde{\mathcal{K}}^{(1)}_{2,2} - v_1\tilde{\mathcal{K}}^{(1)}_{3,2} & \tilde{\mathcal{K}}^{(1)}_{2,3} - v_1\tilde{\mathcal{K}}^{(1)}_{3,3} \\ \tilde{\mathcal{K}}^{(2)}_{1,1} - u_2\tilde{\mathcal{K}}^{(1)}_{3,1} & \tilde{\mathcal{K}}^{(2)}_{1,2} - u_2\tilde{\mathcal{K}}^{(1)}_{3,2} & \tilde{\mathcal{K}}^{(2)}_{1,3} - u_2\tilde{\mathcal{K}}^{(2)}_{3,3} \\ \tilde{\mathcal{K}}^{(2)}_{2,1} - v_2\tilde{\mathcal{K}}^{(2)}_{3,1} & \tilde{\mathcal{K}}^{(2)}_{2,2} - v_2\tilde{\mathcal{K}}^{(2)}_{3,2} & \tilde{\mathcal{K}}^{(2)}_{2,3} - v_2\tilde{\mathcal{K}}^{(2)}_{3,3} \end{pmatrix}$$

$$\text{and } d = \begin{pmatrix} u_1\tilde{\mathcal{K}}^{(1)}_{3,4} - \tilde{\mathcal{K}}^{(1)}_{1,4} \\ v_1\tilde{\mathcal{K}}^{(1)}_{3,4} - \tilde{\mathcal{K}}^{(1)}_{2,4} \\ u_2\tilde{\mathcal{K}}^{(2)}_{3,4} - \tilde{\mathcal{K}}^{(2)}_{1,4} \\ v_2\tilde{\mathcal{K}}^{(2)}_{3,4} - \tilde{\mathcal{K}}^{(2)}_{2,4} \end{pmatrix} .$$

Since we have 4 equations and 3 unknowns we solve

$$X = \underset{Y \in \mathbb{R}^3}{\operatorname{argmin}} \|AY - d\|^2$$

by $X = (A^tA)^{-1}A^td$ (if A^tA is regular). The solution of this linear method can be used as an initial guess for a nonlinear optimization problem. We minimize the distance between the observed points and the projection of the estimated point in the image plane with respect to the image coordinate system: Set $(\tilde{u}_1, \tilde{v}_1, \tilde{w}_1) = \tilde{\mathcal{K}}^{(1)}X$ and $(\tilde{u}_2, \tilde{v}_2, \tilde{w}_2) = \tilde{\mathcal{K}}^{(2)}X$. Then X should minimize

$$\left(u_1 - \frac{\tilde{u}_1}{\tilde{w}_1}\right)^2 + \left(v_1 - \frac{\tilde{v}_1}{\tilde{w}_1}\right)^2 + \left(u_2 - \frac{\tilde{u}_2}{\tilde{w}_2}\right)^2 + \left(v_2 - \frac{\tilde{v}_2}{\tilde{w}_2}\right)^2$$

(see [Zha96]). The solution of this non-linear minimization problem is called the *projective reconstruction* of X.

4.2 Re-projective reconstruction

In the case of the re-projective reconstruction we determine the point of least Euclidean distance to the projection rays. One way to represent a straight line in $\mathbb{R}^3$ is to consider it as an intersection of two planes which are orthogonal to each other. The advantage of this way to model a line, is that the squared Euclidean distance of a point to the line is the sum of squared distances to the planes. These distances can be computed efficiently by the normal forms of the planes. So for the re-projective reconstruction of a point we have to look for a point with the least Euclidean distance to a number of planes.

For $i \in \{1, \ldots, N\}$ let $n^{(i)} = (n_x^{(i)}, n_y^{(i)}, n_z^{(i)})^t$ be the normal and $d^{(i)}$ be the distance of the plane so that $n^{(i)} \bullet \cdot + d^{(i)}$ is the normal form of the plane according to Hesse. The point of the least sum of squared Euclidean distances to all planes is the minimum of

$$\varphi : \begin{array}{l} \mathbb{R}^3 \to \mathbb{R} \\ X \mapsto \sum_{i=1}^{N} \left(n^{(i)} \bullet X + d^{(i)}\right)^2 \end{array} .$$

The minimum of φ can be easily obtained by differentiating φ

$$\varphi'(X) = \sum_{i=1}^{N} 2\left(n^{(i)} \bullet X + d^{(i)}\right) n^{(i)}$$

and finding a solution for $\varphi'(X) = 0$. Since $X = (x, y, z)$ appears only linear in φ' we obtain the result by solving $BX = D$ with

$$B = \sum_{i=1}^{N} n^{(i)} \left(n^{(i)}\right)^t = \begin{pmatrix} \sum_{i=1}^{N} \left(n_x^{(i)}\right)^2 & \sum_{i=1}^{N} n_x^{(i)} n_y^{(i)} & \sum_{i=1}^{N} n_x^{(i)} n_z^{(i)} \\ \sum_{i=1}^{N} n_y^{(i)} n_x^{(i)} & \sum_{i=1}^{N} \left(n_y^{(i)}\right)^2 & \sum_{i=1}^{N} n_y^{(i)} n_z^{(i)} \\ \sum_{i=1}^{N} n_z^{(i)} n_x^{(i)} & \sum_{i=1}^{N} n_z^{(i)} n_y^{(i)} & \sum_{i=1}^{N} \left(n_z^{(i)}\right)^2 \end{pmatrix}$$

and

$$D = -\sum_{i=1}^{N} d^{(i)} n^{(i)} = -\begin{pmatrix} \sum_{i=1}^{N} d^{(i)} n_x^{(i)} \\ \sum_{i=1}^{N} d^{(i)} n_y^{(i)} \\ \sum_{i=1}^{N} d^{(i)} n_z^{(i)} \end{pmatrix} .$$

It is remarkable that in the case of the Euclidean reconstruction of a point no non-linear optimization is needed. Thus the additional framework to obtain the suitable representation of the projection rays as an intersection of two orthogonal planes in normal form is justified.

5 Experimental results

5.1 8.5 mm lens stereo setup

One test to compare both calibration methods is the 3D-reconstruction of points by a stereo camera system. The setup we used consists out of two common CCD cameras with 8.5 mm lenses. To calibrate our cameras we use 8 images of a calibration pattern with a grid of 20×19 dots with 20 cm distance to each other (see Fig. 2).

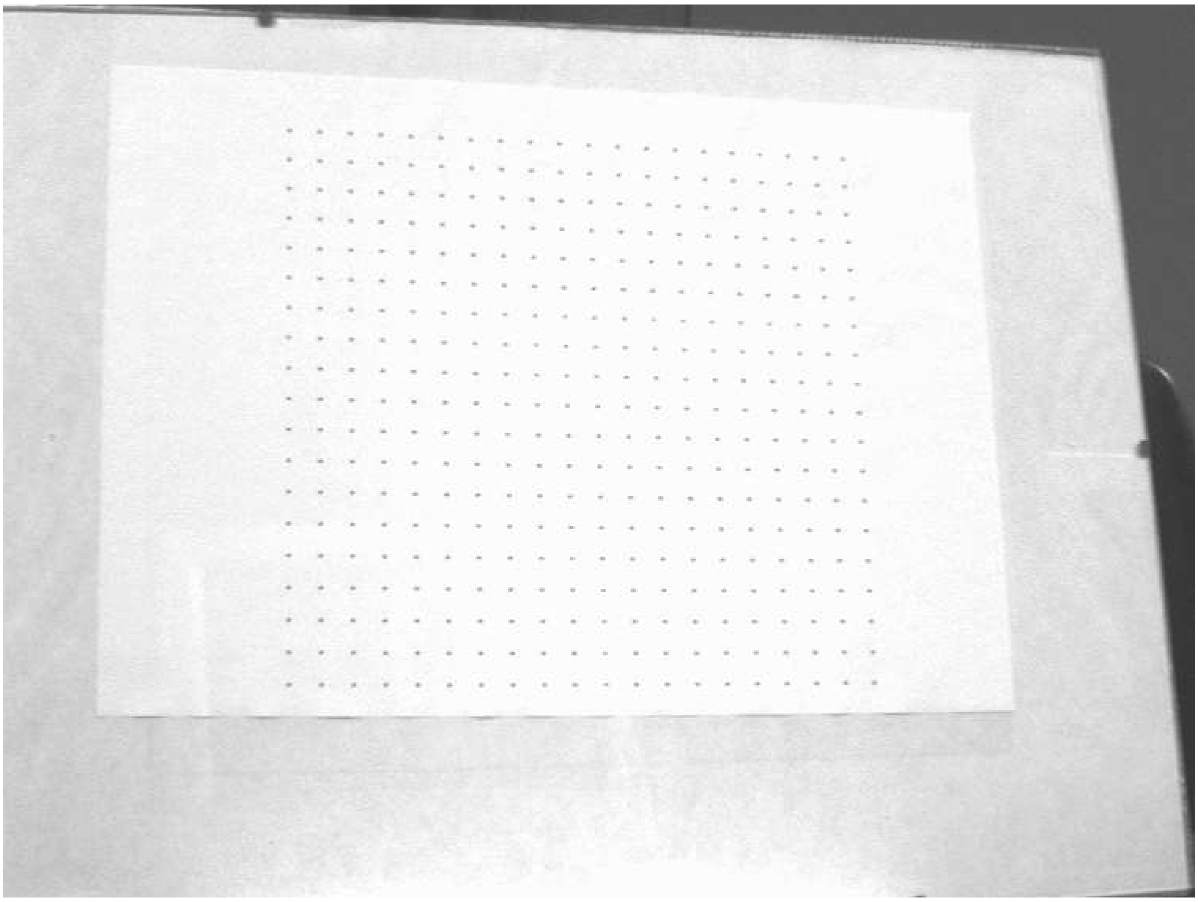


Figure 2: Calibration grid of 20×19 points

In Table 1 we list the resulting values of both error functions (Eq. 2 and Eq. 3). Obviously in the case of the re-projective calibration the Euclidean error is lower than the same value for the projective calibration since this is the objective of this error function. Consequently the projective error for the projective calibration method is lower than the same value for the re-projective calibration. The resulting camera parameters for both methods and cameras are listed in table 2.

To compare the two calibration methods we define two 3D-reconstruction problems:

Table 1: Minimal values of the optimized error functions – projective and re-projective – for both calibration methods and both cameras

	Error left camera		Error right camera	
Calibration	Projective	Re-proj.	Projective	Re-proj.
Projective	186.894 pix	0.325452 m	172.779 pix	0.307025 m
Re-proj.	186.983 pix	0.325220 m	173.151 pix	0.305378 m

Table 2: Obtained camera parameters for projective and re-projective calibration (8.5 mm lenses)

	Left camera		Right camera	
	Projective	Re-projective	Projective	Re-projective
α	1301.082	1301.057	1359.023	1359.887
β	1305.719	1305.638	1365.126	1366.030
γ	0.783055	0.699502	1.182897	1.464901
u_0	345.5133	345.5417	363.0120	362.6659
v_0	280.0936	280.5081	272.2942	271.1639
k_1	-0.12926	-0.12746	-0.13097	-0.12594
k_2	-0.41933	-0.45143	-0.07319	-0.12728

we measure 3D-distance of points reconstructed from our calibration pattern and the planarity of the reconstructed 3D-plane.

Reconstruction test

For the first comparison we reconstruct each point of the calibration pattern. To measure the result we compare the distance of the reconstructed points to their neighbors. This distance is 20 mm on the calibration pattern and therefore should be 20 mm in the reconstruction. In Table 3 we denote the average distance between neighbored reconstructed points less 20 mm. The columns "Projective" and "Re-proj." denote the calibration algorithm. For each image we performed the projective and the re-projective reconstruction of the points. We took four image pairs at different positions and angles to the cameras. These images were not used for the calibration itself.

Except image pair 3, the re-projective calibration with the re-projective reconstruction method yield the smallest difference to the true distance between the points. Using only the projective calibration method, the re-projective reconstruction method almost gives the best results. Most of the image pairs also have better results for the re-

Table 3: Average of distances between neighbored points in mm (less 20 mm) for both calibration and reconstruction methods

Image pair	Reconstruction method	Calibration method	
		Projective	Re-proj.
1	Projective	-0.00252	0.00106
	Re-projective	-0.06062	0.00089
2	Projective	0.00340	-0.03319
	Re-projective	-0.04935	0.00304
3	Projective	0.00888	0.00892
	Re-projective	0.00708	0.00908
4	Projective	0.06082	0.00370
	Re-projective	0.00888	0.00368

projective calibration when only considering the projective reconstruction method. The re-projective methods (calibration and reconstruction) are advantageous for this test.

Planarity test

For the planarity test we fit a plane through the reconstructed points and measure the distance of the points to the fitted plane. The results in Table 4 are obtained for the same set of images as for the reconstruction test. We compare the result for the projective (column “Projective”) and the re-projective (column “Re-proj.”) calibration algorithm. Again we distinguish between projective and re-projective reconstruction of the points.

Table 4: Average of distance in mm to the fitted plane for both calibration and reconstruction methods

Image pair	Reconstruction method	Calibration method	
		Projective	Re-proj.
1	Projective	5.239189	5.425621
	Re-projective	5.307714	5.454959
2	Projective	3.556000	3.604187
	Re-projective	3.425454	3.631118
3	Projective	5.334670	5.403838
	Re-projective	5.129301	5.436539
4	Projective	5.638853	5.695013
	Re-projective	5.393518	5.732125

In this case the projective methods – calibration and reconstruction – naturally take advantage of the character of projective mappings: planes are mapped on planes. Therefore the first two image pairs yield the best results for this combination of methods. Nevertheless, the smallest error for the last two image pairs is reached when using the re-projective reconstruction method. Note that a small distance to the fitted plane implies not that the reconstructed (fitted) plane has a small distance to the original plane.

5.2 4.8 mm monocular setup

For another experiment we choose one camera with 4.8 mm lens with significant distortion. To calibrate the camera we choose a 9×5 calibration plate at a distance of about 1 m (see Fig. 3). For this setup the results of the projective and re-projective calibration algorithm differ more than those for the 8.5 mm lenses (see Table 5). It is remarkable that the re-projective calibration algorithm improves the Euclidean error by factor 1000 but leaves the projective error nearly untouched (see Table 6). In fact the projective calibration turned out to be very difficult: The classic approach of Zhang [Zha98] (closed form calibration followed by nonlinear optimization without the distortion parameters before really applying Eq. 2) leads to a much worse local minimum. We had to apply the nonlinear optimization of Eq. 2 directly after the closed form solution.

Table 5: Obtained camera parameters for projective and re-projective calibration (4.8 mm lens)

	Projective	Re-projective
α	766.3934	766.4256
β	766.7465	767.0491
γ	2.798895	2.656353
u_0	379.4183	379.5317
v_0	283.5199	282.0707
k_1	-0.26054	-0.26748
k_2	-0.12400	-0.14375

Monocular reconstruction of a calibration pattern

Obviously the monocular setup allows no stereo reconstruction of single points. For a rotation R and a translation t we denote $T_{R,t}$ as the resulting transformation of points. Since we know the position of each point on the plate we can obtain the position of plate by minimizing

$$(R,t) \mapsto \sum_{p \in \mathbf{P}} \|i_p - \Pi \circ T_{R,t}(p)\|^2$$

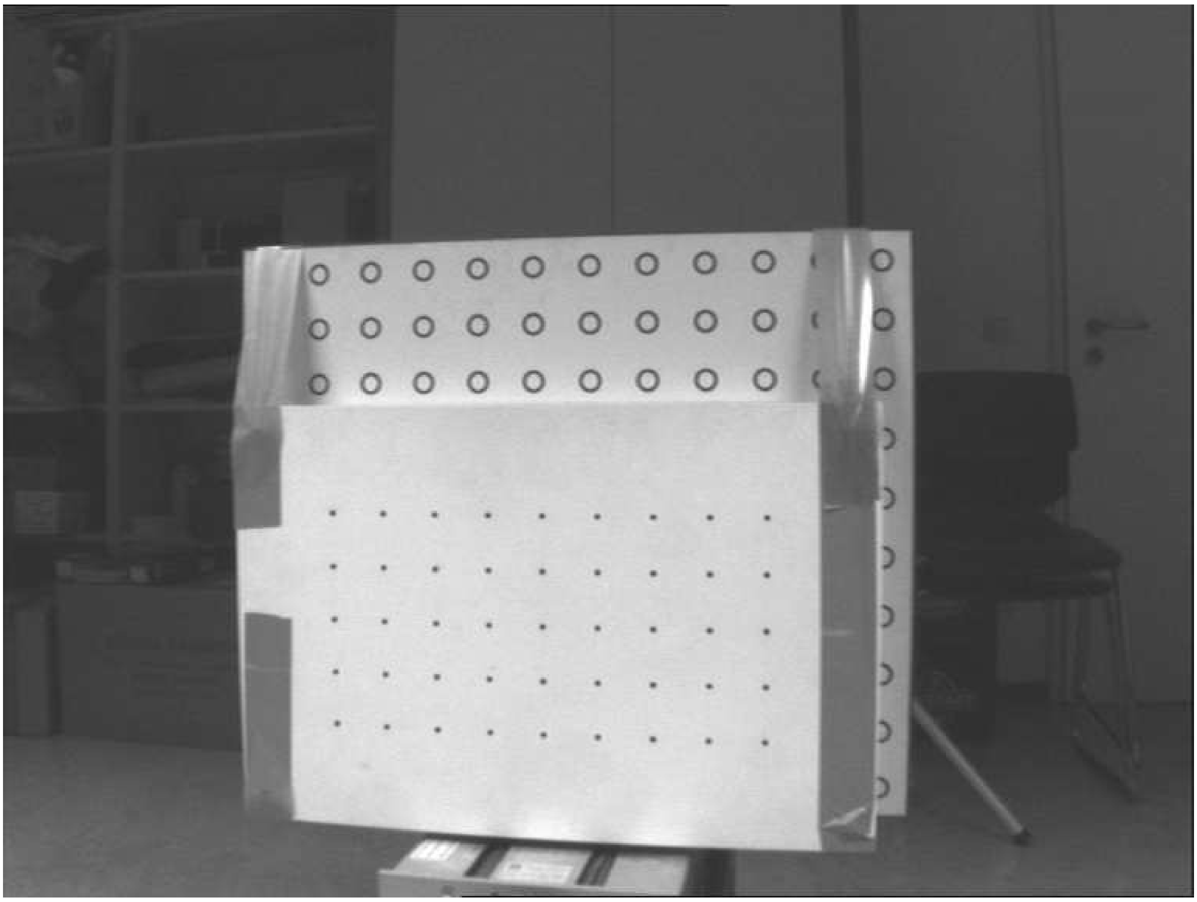


Figure 3: Calibration plate (9×5 points) on slide

Table 6: Minimal values of the optimized error functions for 4.8 mm lens

Calibration method	Error function	
	Projective	Re-projective
Projective	24.513411 pix	72.73676 mm
Re-projective	24.625616 pix	0.08408 mm

in the projective case and

$$(R, t) \mapsto \sum_{p \in \mathbf{P}} \mathrm{dist}_3(T_{R,t}(p) - \Pi^{-1}(\{i_p\})^2$$

in the re-projective case (see [HP99] for details of the re-projective approach). In any case the non-linear optimization is performed by using the Levenberg-Marquardt-algorithm.

Moving plane test

We fixed our calibration pattern on a linear sledge which can be moved on one axis with a very high precision (<0.01 mm). Therefore we are able to measure the exactness of a monocular reconstruction of the calibration with respect to the zero position of the sledge: We move the plate into the background with a distance of 3 cm between each observed image. After the sixth position we move the plate back to the foreground. In Table 7 a comparison of the two calibration algorithms combined with the two reconstruction methods is shown. For the 7th position in both parts of the table we did not move the plane, but we took two different images. So the results in these rows (displayed bold face) can be seen as the noise of the signal.

For all positions of the calibration plate the minimal error is reached using a reprojective method (calibration or reconstruction), but it must be admitted that the improvement is of a small magnitude.

Table 7: Results of the moving plate test for the 4.8 mm lens in cm

Pos. of sledge	Recon-struction method	Calibration method			
		Projective		Re-projective	
		Position	Error	Position	Error
1	Projective	2.9583	0.0417	2.9592	0.0408
	Re-proj.	2.9299	0.0701	2.9591	0.0409
2	Projective	2.9804	0.0196	2.9812	0.0188
	Re-proj.	2.9538	0.0046	2.9810	0.0190
3	Projective	2.9376	0.0624	2.9383	0.0617
	Re-proj.	2.9137	0.0086	2.9382	0.0618
4	Projective	2.9762	0.0238	2.9770	0.0230
	Re-proj.	2.9514	0.0049	2.9769	0.0231
5	Projective	2.9800	0.0200	2.9809	0.0191
	Re-proj.	2.9551	0.0045	2.9809	0.0191
6	Projective	2.9649	0.0351	2.9654	0.0346
	Re-proj.	2.9430	0.0570	2.9652	0.0348
7	Projective	**0.0143**		**0.0144**	
	Re-proj.	**0.0130**		**0.0144**	
8	Projective	2.9567	0.0433	2.9575	0.0425
	Re-proj.	2.9300	0.0700	2.9573	0.0427
9	Projective	2.9732	0.0268	2.9737	0.0263
	Re-proj.	2.9528	0.0472	2.9738	0.0262
10	Projective	2.9500	0.0500	2.9510	0.0490
	Re-proj.	2.9239	0.0761	2.9508	0.0492
11	Projective	2.9625	0.0375	2.9631	0.0369
	Re-proj.	2.9376	0.0624	2.9630	0.0370
12	Projective	2.9798	0.0202	2.9808	0.0192
	Re-proj.	2.9513	0.0487	2.9807	0.0193
13	Projective	2.9656	0.0344	2.9661	0.0339
	Re-proj.	2.9416	0.0584	2.9661	0.0339

6 Conclusion and discussion

In this article we presented two different objective functions for the camera calibration problem. The well known projective error function minimizes the distance of projected calibration points to observed points. The re-projective error function points out the re-projective property of a camera: Each observed point in the image plane defines a straight line in the reference coordinate system, the re-projected line. The resulting error function minimizes the distance of the calibration points to the re-projected lines.

This error function for camera calibration minimizes the error which should also be minimized in 3D-reconstruction. Since our reference coordinate system is Euclidean, we should also measure the distance in the Euclidean sense. This means that we should consider the distance of points to re-projected lines as orthogonal projection problem and not as a 2D-distance in a projective space.

The experimental results in section 5.1 encourage to do more research in the field of re-projective camera calibration.

References

[DF95] F. Devernay and O. Faugeras. Automatic calibration and removal of distortion from scenes of structured environments. In *SPIE Conference on investigative and trial image processing*, SanDiego, CA, 1995.

[Hec87] Eugene Hecht. *Optics*. Addison-Wesley, Reading, Ma, 2. edition, 1987.

[HGP04] T. Hanning, S. Graf, and G. Pisinger. Extrinsic calibration of a stereo camera system fulfilling generalized epipolar constraints. In *International Conference on Visualization, Imaging and Image Processing Conference (VIIP)*, pages 1–5, Marbella, Spain, 2004.

[HP99] T. Hanning and G. M. Pisinger. 3d reconstruction from monucular image sequences. In *Diagnostic image technologies and applications*, EurOpto, pages 121–132, Munich, 1999.

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

[Mar63] D. W. Marquardt. An algorithm of least-squares estimation of nonlinear parameters. *J. Soc. Indust. Appl. Math.*, 11:431 – 441, 1963.

[PWH97] T. Pajdla, T. Werner, and V. Hlaváč. Correcting radial lens distortion without knowledge of 3-d structure. Technical Report K335-CMP-1997-138, Czech Technical University, Center for Machine Perception, Praha, 1997.

[Tsa87] Roger 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.

[TYO02] T. Tamaki, T. Yamamura, and N. Ohnishi. Unified approach to image distortion. In *International Conference on Pattern Recognition*, volume II, pages 584–587, Quebec, Canada, August 2002.

[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.

[Zha96] Z. Zhang. Determining the epipolar geometry and its uncertainty: A review. Technical Report 2927, INRIA, 1996.

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