Affine Stereo Calibration

Peter Sturm, Long Quan

Affine stereo calibration

Peter Sturm and Long Quan
GRAVIR-IMAG & INRIA Rhône-Alpes*
655, Avenue de l'Europe, 38330 Montbonnot, France
Peter.Sturm@inrialpes.fr Long.Quan@inrialpes.fr

26/09/1995

Abstract

Affine stereo calibration has been identified as the determination of the plane collineation induced by the plane at infinity. Generally, this so-called infinity homography can be determined provided 3 image correspondences of points at infinity.

In this paper, we are first concerned with affine stereo calibration for the case of two cameras with the same intrinsic parameters. It is shown that the affine calibration is possible with one less point at infinity due to the invariance of the intrinsic parameters for the two cameras.

Then, we propose two practical methods for affine stereo calibration in the general case. The experimental results both on simulated and real images are presented. The quality of the calibration method is evaluated by the quality of affine reconstruction.

Key words: affine calibration, affine reconstruction, invariant, geometry.

1 Introduction

A basic task in computer vision applications is the determination of camera parameters, the camera calibration. Camera calibration is the basis of how to obtain 3D reconstructions which can be used to recognize objects, navigate in an unknown environment, or just to give binary informations about the scene, like "there is/is not an obstacle in front of the cameras". It is evident that, for all three examples, exact 3D reconstruction is not always necessary. Object recognition can be realized on the basis of Euclidean as well as affine or projective invariants. Navigation sub-tasks often require knowledge of only affine instead of Euclidean scene properties. For simple obstacle detection the identification of a plane in the scene and the epipolar geometry of the camera system are sufficient.

With regard to these examples and due to the non-stability of calibration processes, it is interesting to examine the possibilities offered by cameras that are not completely calibrated. With the aim of 3D reconstruction in mind, one can identify the three principal levels of stereo calibration [2, 6, 10, 8, 4, 13]: Eucidean, affine, and projective calibration. The notations are chosen in order to reflex the nature of 3D reconstructions rendered possible on the different calibration levels: 3D reconstruction up to a Euclidean, affine, or projective transformation.

In the pioneer work of Faugeras [2] on projective reconstruction from two uncalibrated images, the affine reconstruction proposed was determined up to three free parameters. This is later clarified in [10] in which it was shown that the affine calibration is equivalent to determine the infinity homography between the two image planes.

The goals of this paper are twofold. Firstly, we explore the theory of affine calibration for cameras with same intrinsic parameters, or, equivalently, for one camera in movement. Secondly we emphasize practical aspects of affine calibration which lead to the proposition of two practical methods of affine calibration.

The paper is organized as follows. In section 2 we describe the camera model used troughout the paper and several other geometric preliminaries. Section 3 gives a brief overview of different concepts of partial stereo calibration and affine reconstruction. Section 4 is dedicated to the detection of vanishing points and lines which is the basis for the implementations of our affine calibration methods. Then, in section 5, we investigate the problem

*This work has been done in the context of the Movi project which belongs to Cnrs, Inpg, Inria and Ujf.

of affine calibration in the special case of two cameras having the same intrinsic parameters. In section 6 we present two practical methods of affine stereo calibration. For one of the calibration methods, the results obtained in experiments with synthetic as well as real data are presented in section 7. Section 8 illustrates the usefulness of affine calibration by outlining a simple simulation of a vehicle which is navigating with the aid of an affinely calibrated stereo system. Finally, in section 9, we conclude this paper with a short discussion. In four appendices we describe some methods for the determination of vanishing points and lines.

2 Some preliminaries

2.1 Camera model

The camera model used throughout the paper is the pin-hole model which models the camera projection as a perspective transformation from $\mathcal{P}^3$ to $\mathcal{P}^2$, represented by a matrix $P_{3\times4}$. This camera or projection matrix can be decomposed as

$$P_{3\times4} = KD = \underbrace{\begin{pmatrix} \alpha_u & -\alpha_u \cot\Theta & u_0 \\ 0 & \alpha_v/\sin\Theta & v_0 \\ 0 & 0 & 1 \end{pmatrix}}_{K} \underbrace{\left(\begin{array}{c|c} R_{3\times3} & \mathbf{t}_3 \end{array}\right)}_{D} = \left(\begin{array}{c|c} KR & K\mathbf{t} \end{array}\right) = \left(\begin{array}{c|c} \bar{P}_{3\times3} & \mathbf{p}_3 \end{array}\right)$$

where R and $\mathbf{t}$ represent the orientation and position of the camera and K its projection properties. The 6 parameters of R and $\mathbf{t}$ are called extrinsic camera parameters, K is determined by the 5 intrinsic camera parameters.

The pin-hole camera can be illustrated by two elements, the *image plane* and the *center of projection*. The projection of any 3D point is then just the intersection of image plane and the line joining the center of projection and the 3D point.

2.2 Plane collineation

The pin-hole model just described does a perspective transformation from $\mathcal{P}^3$ to the image plane. This transformation induces a one-to-one transformation between any plane Π in $\mathcal{P}^3$, that does not contain the center of projection, and the image plane. Let H_1 be this transformation. When considering a second camera, there exists analogously a transformation H_2 between Π and the image plane of this second camera.

Let Q be any point on Π and $\mathbf{q}$ and $\mathbf{q}'$ its projections by the two cameras. We have the following relationship[1]:

$$\mathbf{q}' \approx \underbrace{H_2 H_1^{-1}}_{H_\Pi} \mathbf{q}$$

The transformation H_Π is called *plane collineation, induced by the plane* Π. It can be represented by a 3×3 matrix which is defined up to a scalar factor.

2.3 Epipolar geometry

The projection of the center of projection of one camera by a second camera is called the *epipole* of the second camera. We denote the epipoles of a system of two cameras as $\mathbf{e}$ and $\mathbf{e}'$ (see Fig 1).

Consider a 3D point Q and its projections $\mathbf{q}$ and $\mathbf{q}'$. Q and the centers of projection, C and C', are spanning a plane, which is called *epipolar plane* of Q, since it contains both of the epipoles. This epipolar plane intersects the first image plane R in the line $\langle \mathbf{e}, \mathbf{q} \rangle$ and the second one in $\langle \mathbf{e}', \mathbf{q}' \rangle$. We call the respective line in the i*th* image plane i*th epipolar line* of Q.

The set of all epipolar lines in one image plane forms a line pencil, with the epipole as node. The *epipolar transformation* is defined as the projective transformation from the first to the second epipolar pencil, which is induced by the intersection of the image planes with the epipolar planes. The two epipoles together with the epipolar transformation define entirely the *epipolar geometry*, some aspects of which we discuss in the following.

[1] By $\approx$ we mean equaltity up to a scalar factor

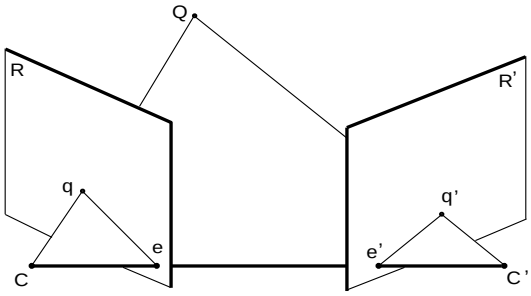


Figure 1: Stereosystem and epipolar geometry

Consider a point $\mathbf{q}$ in the first image plane, which is the projection of an unknown 3D point Q. The epipolar plane, which is spanned by $\mathbf{q}$ and the centers of projection, must contain Q and thus also its second projection $\mathbf{q}'$. So, $\mathbf{q}'$ must lie on the epipolar line, which corresponds to $\langle \mathbf{e}, \mathbf{q} \rangle$ via the epipolar transformation. This is the so-called *epipolar constraint*.

The epipolar geometry can be represented by the *Fundamental Matrix* F, a 3×3 matrix of rank 2 [2]. The epipolar constraint is expressed by the condition

$$(\mathbf{q}')^T F \mathbf{q} = 0. \tag{1}$$

2.4 Double ratio

The *double ratio* is the basic projective invariant. It can be expressed in different forms, where the simplest one is the double ratio of four collinear points:

$$\text{double ratio}_{Q,R,S,T} = \{Q, R; S, T\} = \frac{|SQ|}{|SR|} / \frac{|TQ|}{|TR|}$$

where $|.|$ is the Euclidean distance between points. Here, the distance between a point at infinity and an affine point is considered as infinity. Furthermore, the following rules are applicated:

$$\frac{0}{0} = 1 \qquad \frac{\infty}{\infty} = 1 \qquad \infty * 0 = 1$$

2.5 Absolute conic and co-circular points

Circles in the affine plane always intersect in the two *co-circular points* I and J, which are given by [12]

$$I = \begin{pmatrix} 1 \\ +i \\ 0 \end{pmatrix}, \; J = \begin{pmatrix} 1 \\ -i \\ 0 \end{pmatrix}.$$

I and J are imaginary points at infinity which are complex conjugated. The co-circular points of all planes in $\mathcal{P}^3$ form a 3D conic section, which is called *absolute conic*.

3 A brief overview of stratification of partial stereo calibration

3.1 Projective calibration

Projective calibration[2] consists on the determination of the epipolar geometry of two cameras [2]. The epipolar geometry is entirely representated by the fundamental matrix $F_{3\times3}$ which is of rank 2. The fundamental matrix enables the determination of the projection matrices of the two cameras, up to an unknown, but common, projective transformation [5].

[2]often also called weak calibration [2]

3.2 Quasi-affine calibration

Projective calibration enables projective 3D reconstruction of the scene which is very weak, for example it is not oriented. If we, for a given scene, know that the plane at infinity does not cross the scene, then we are able to determine 3D placement of points with respect to reference planes. In this context, quasi-affine calibration is, in addition to a projective calibration, the identification of a plane Π in the two images, i.e. the determination of the plane collineation induced by Π. In this case we are able to determine the relative placement of 3D points to Π which means that we can determine if a point lies on Π, behind or before Π, viewed from the camera system. This calibration step can for instance be applicated for obstacle detection with respect to a ground plane.

The plane collineation can be determined from 3 correspondences for projections of points on Π (the $4th$ point correspondence which is needed to fix the 8 degrees of freedom of the plane collineation, is always given by the two epipoles).

3.3 Affine calibration

Affine calibration is the special case of quasi-affine calibration, when it is the plane at infinity whose plane collineation is determined. In the sequel, we call this special plane collineation *infinity homography* and represent it by the 3×3 matrix H_∞. The infinity homography can be computed by [10]

$$H_\infty \approx \bar{P}'(\bar{P})^{-1} = K'R'(KR)^{-1} = K'R'R^TK^{-1}. \tag{2}$$

The infinity homography can be determined from 3 correspondences of vanishing points.

Affine calibration enables affine reconstruction of the scene, for example by the two methods described in the following paragraphs. Affine invariants of a scene, like for example the mid-point of two points or the parallel of a given line, passing through a given point, can be found, given an affine calibration. The ability of determining mid-points is used in the simulation described in section 8.

3.3.1 Affine reconstruction using reference points.

It is a current method to reconstruct a scene using reference points in the images which define an affine or projective basis [7, 2, 9, 10]. In the case of affine reconstruction, we can select any four points (no three of them collinear) and assign them the canonical affine basis:

$$\mathbf{b}_1 = (0,0,0)^T, \mathbf{b}_2 = (1,0,0)^T, \mathbf{b}_3 = (0,1,0)^T \text{ , and } \mathbf{b}_4 = (0,0,1)^T.$$

Then, if $\mathbf{x} = \mathbf{b}_2 - \mathbf{b}_1$, $\mathbf{y} = \mathbf{b}_3 - \mathbf{b}_1$, and $\mathbf{z} = \mathbf{b}_4 - \mathbf{b}_1$, any fifth point's coordinates can be expressed in the form $\alpha\mathbf{x} + \beta\mathbf{y} + \gamma\mathbf{z}$.

$(\alpha, \beta, \gamma)^T$ are the desired affine coordinates which can be uniquely determined in the images through appropriate geometric operations [10] which take the infinity homography into account.

3.3.2 Reconstruction using a reference view.

Another reconstruction approach is to take one of the images as the reference frame [5, 8] instead of selecting reference points. In the context of affine reconstruction it has been shown [8] that the matrices[3]

$$\left(\begin{array}{c|c} I_3 & \mathbf{0}_3 \end{array}\right) \quad \text{and} \quad \left(\begin{array}{c|c} H_\infty & \mathbf{e}' \end{array}\right)$$

equal the projection matrices, up to an affine transformation in $\mathcal{P}^3$. Consequently, if we use these matrices for reconstruction of a scene, for example by pointwise triangulation [3], we obtain a 3D representation up to an (unknown) affine transformation.

[3] $\mathbf{e}'$ is the epipole in the second image plane

4 Detection of vanishing points and lines

Affine calibration, as it consists in the determination of the infinity homography, requires the detection of image correspondences for structures at infinity. In this section we describe one method for detection of vanishing points and one for vanishing lines. These are used by the affine calibration methods proposed in section 6.

Some other methods for detection of vanishing points or lines, which are of a more theoretic interest, are described in the appendix.

4.1 Detection of vanishing points

We consider 2 parallel lines in space. The projection of their common point at infinity, their vanishing point, is obtained by intersecting the projections of the lines in the image plane. Usually, if we consider more than 2 lines, this becomes a fitting problem which requires a minimization technique.

4.2 Detection of vanishing lines

We consider 2 parallel planes and describe in the following how to determine the projections of their intersection line, which are vanishing lines (another method is described in section B). We consider the minimal case, i.e. 2 planes, each one spanned by 3 points, which are projected into the two image planes (look Fig 2 (a)). It is supposed to already dispose of a projective calibration of the stereo system and that the matching between projections of the 2×3 points is done.

First, we carry out a projective reconstruction of the 2 planes, using the projective calibration which is given. In order to do this we reconstruct (projectively) the 2×3 points [3] (Fig 2 (b)) and obtain the planes in the projective space by spanning the 2 sets of respectively 3 points (Fig 2 (c)). Then, we intersect the planes and project the intersection line into the image planes, using the projective calibration information (Fig 2 (d)). The projected lines are the vanishing lines of the 2 planes, since intersection of planes is a projective invariant.

If more than 2 planes with eventually more than 3 points are given, minimization techniques have to be applied.

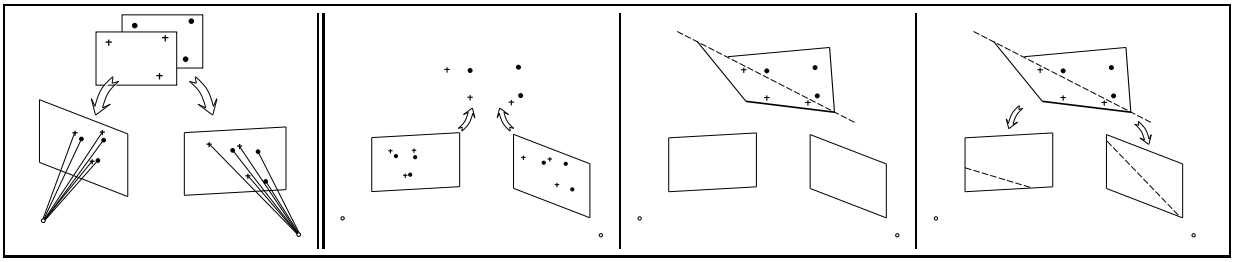


Figure 2: Determination of the vanishing line of two parallel planes.
(a) The projections of points of interest on parallel planes are the required data. **(b)** Projective reconstruction of the 3D points. **(c)** Spanning the planes in projective space. **(d)** The projections of the intersection line are the searched vanishing lines.

5 Affine calibration of two cameras with same intrinsic parameters

We consider the case of two cameras having same intrinsic parameters, or, equivalently, of one camera which is moved once. In this case, a constraint on the infinity homography arises which should be taken into account when estimating H_∞ [8]. Reciproquely, this constraint leads to that H_∞ can be determined, up to 6 solutions, from only 2 correspondences of vanishing points, instead of 3 in the general case.

5.1 A constraint for the infinity homography

The two cameras having the same intrinsic parameters, i.e. $K' = K$, changes equation 2 into

$$H_\infty \approx KR'(KR)^{-1} = KR'R^TK^{-1} = KR''K^{-1}$$

where R'' is a rotation matrix. This equation means nothing else, that H_∞ is similar to a multiple[4] of R''. Hence, H_∞ has eigenvalues $\lambda, \lambda e^{i\theta}$, and $\lambda e^{-i\theta}$, with a $\lambda \in \mathcal{R}$. Thus, the eigenvalues of H_∞ have the same module. This constraint should be taken into account in any estimation of H_∞, where the two cameras have same intrinsic parameters, as has already been stated in [8].

5.2 Determination of the infinity homography from two correspondences of vanishing points

In this paragraph we give a proof for the following proposition.

Proposition 1. *Suppose that two cameras have same intrinsic parameters and the fundamental matrix for the cameras is known. Suppose further that the correspondences of 2 vanishing points are known.*

Then, in non-degenerated cases [5]*, we can determine the infinity homography up to at most 6 solutions.*

Proof. Since we already dispose of 2 vanishing point correspondences, our aim is to construct a third one which will enable to calculate H_∞. Let us denote the known vanishing points as $\mathbf{v}$ and $\mathbf{w}$ for the first image, respectively as $\mathbf{v}'$ and $\mathbf{w}'$ for the second one. The third point correspondence we are searching for are the points $\hat{\mathbf{x}}$ and $\hat{\mathbf{x}'}$.

Our strategy is to take an arbitrary point in the first image as $\hat{\mathbf{x}}$, with the only condition that it is neither aligned with $\mathbf{v}$ and $\mathbf{w}$, nor is the epipole $\mathbf{e}$. Then, we pursue the corresponding epipolar line in the second image, and, for each point $\mathbf{x}'$ on it, we calculate the plane collineation defined by the point correspondences $(\mathbf{v} \leftrightarrow \mathbf{v}')$, $(\mathbf{w} \leftrightarrow \mathbf{w}')$, $(\hat{\mathbf{x}} \leftrightarrow \mathbf{x}')$, and $(\mathbf{e} \leftrightarrow \mathbf{e}')$. Let us denote this plane collineation by $H(\mathbf{x}')$, since $\mathbf{x}'$ is the only changing parameter.

We now pick out those $H(\mathbf{x}')$ which are similar to a multiple of a rotation matrix, i.e. whose eigenvalues all have the same module. In the following we show that there are at most 6 $H(\mathbf{x}')$ which have eigenvalues with all the same module.

The plane collineation H_Π for an arbitrary plane Π is of the form [8]

$$H = K(R'' + \frac{1}{d}\mathbf{t}\mathbf{n}^T)K^{-1} \tag{3}$$

where R'' and $\mathbf{t}$ are the rotation and translation which lie between the two cameras, and $(\mathbf{n}_3, d)$ is the representation of Π by normal vector and distance from the origin.

Since the calculation of any $H(\mathbf{x}')$ includes the correspondences of two same vanishing point correspondences, $(\mathbf{v} \leftrightarrow \mathbf{v}')$ and $(\mathbf{w} \leftrightarrow \mathbf{w}')$, all thereby referenced planes $\Pi(\mathbf{x}')$ are parallel (they all contain the same line at infinity whose projections are $\langle \mathbf{v}, \mathbf{w} \rangle$ and $\langle \mathbf{v}', \mathbf{w}' \rangle$). Since parallel planes have the same normal vector, the only variable parameter when applying equation 3 for the $H(\mathbf{x}')$, is the scalar d.

We deduce, that $H(\mathbf{x}')$ is similar to the matrix $\mu(R'' + d'M)$, with a $\mu \in \mathcal{R}$, $M = \mathbf{t}\mathbf{n}^T$, and $d' = \frac{1}{d}$. The characteristical polynomial of $H(\mathbf{x}')$,

$$f_{H(\mathbf{x}')} = \tau^3 + c_2(d')\tau^2 + c_1(d')\tau + c_0(d') \tag{4}$$

is a polynomial of degree 3 in d' (c_2 is of degree 1, c_1 of degree 2, and c_0 of degree 3).

Applying the condition, that $H(\hat{\mathbf{x}'})$ has to have eigenvalues $\lambda, \lambda e^{i\theta}$, and $\lambda e^{-i\theta}$, its characteristical polynomial must be of the form

$$f_{H(\hat{\mathbf{x}'})} = (\tau - \lambda)(\tau - \lambda e^{i\theta})(\tau - \lambda e^{-i\theta}) = \tau^3 - \lambda(1 + 2\cos\theta)\tau^2 + \lambda^2(1 + 2\cos\theta)\tau - \lambda^3. \tag{5}$$

Combining equations 4 and 5 leads to the following conditions for $H(\mathbf{x}')$ to be similar to a multiple of a rotation matrix.

$$\begin{aligned} c_2(d') &= -\lambda(1 + 2\cos\theta) \\ c_1(d') &= \lambda^2(1 + 2\cos\theta) \\ c_0(d') &= -\lambda^3. \end{aligned}$$

[4] Multiple of a matrix means here the matrix, multiplied with any scalar

[5] This is precised in the proof

The combination of these equations leads to the single condition

$$c_1(d') = \sqrt[3]{c_0(d')}c_2(d') \ .$$

This is equivalent to

$$g(d') := c_1^3(d') - c_0(d')c_2^3(d') = 0 \tag{6}$$

where $g(d')$ is a polynomial of degree 6 in d'.

Hence, in the non-degenerated case, when not all 7 coefficients of g vanish, g has at most 6 real roots in d'. So, there are at most 6 $H(\mathbf{x}')$ which are similar to a multiple of a rotation matrix. The final conclusion of the proof is, that we can determine $H_\infty \approx H(\hat{\mathbf{x}'})$ up to at most 6 solutions $H(\mathbf{x}')$. QED.

The constructional nature of the proof permits its relatively easy transformation into a practical algorithm.

6 Practical affine calibration methods

We propose two practical methods of affine calibration in the general case, i.e. the cameras need not to have the same intrinsic parameters. Both of the methods suppose a projective calibration being given.

6.1 First affine calibration method: Observation of a plane during a translational movement

We consider a plane Π_1 which contains 3 points of interest which are detected and matched in the two images. The plane Π_1 is subject to a translational movement and is so mapped onto a plane Π_2 parallel to Π_1. We suppose that the images of the 3 points of interest can also be detected for Π_2 and that the correspondence between the images of corresponding points of interest is made. The configuration is shown in Fig 3. We observe that:

- The planes Π_1 and Π_2 are parallel.
- The 3 lines, which join the points of interest on Π_1 with the corresponding ones on Π_2, are parallel.

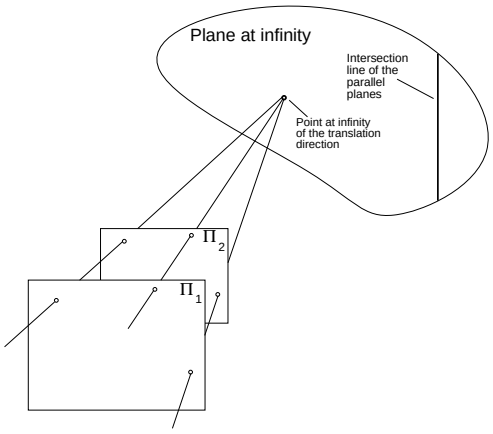


Figure 3: Multiple translation of a plane in the same direction and accompanying structures at infinity.

This knowledge can be used to determine one correspondence of vanishing points and one of vanishing lines with the methods described in section 4. Once these correspondences have been established, we can estimate the infinity homography.

The extension of the method for more than 3 points of interest or more than 2 parallel planes is straightforward in accordance with what is said in section 4.

6.2 Second affine calibration method: Three translational movements

We consider two points in the scene, Q_1 and R_1, and suppose that they are subject to three consecutive translational movements in different directions (see illustration in Fig 4). Let Q_{i+1} and R_{i+1} be the points after the ith translation. We observe that each two lines $\langle Q_i, Q_{i+1}\rangle$ and $\langle R_i, R_{i+1}\rangle$ for $i \in \{1, 2, 3\}$ are parallel. The three pairs of parallel lines allow to detect three correspondences of vanishing points in the image planes (see section 4). Given these correspondences, the infinity homography can be estimated.

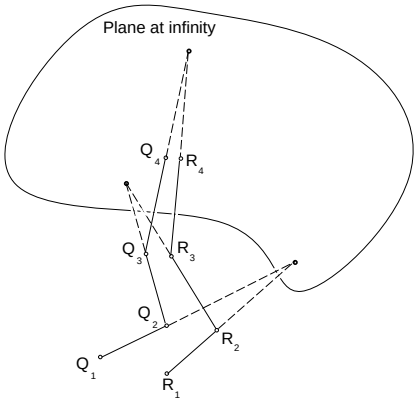


Figure 4: Multiple translation of points in different directions. The points at infinity of the translation directions are shown.

This method can be straightforwardly extended to more than 2 points or more than 3 translational movements.

7 Experimental results

7.1 Outline of the experiments

For the first of the two described affine calibration methods (see section 6.1) we undertook experiments with simulated as well as real image data. The different steps of an experimentation process are:

1. Projective calibration using an iterative algorithm.
2. Affine calibration of the stereo system with the method described in section 6.1.
3. Euclidean calibration with a classical method, using a completely known calibration object.
4. Affine reconstruction of a scene (see paragraph 3.3.2).
5. Euclidean reconstruction of the same scene. In experimentation with simulated data we take here the original 3D point set.
6. Evaluating the quality of the affine reconstruction by comparing it with the Euclidean one (see paragraph 7.2). The obtained quality of affine reconstruction serves as quality mesure for the affine calibration method.

7.2 Evaluation of the quality of affine reconstruction

Suppose that we have reconstructed n points, where the Q_i are the affine reconstruction and the R_i the Euclidean reconstruction. We evaluate the affine reconstruction by the following value:

$$\frac{1}{n} \min_{T \in \mathcal{A}_3} \sum_{i=1}^{n} |R_i - TQ_i| \tag{7}$$

Gaussian noise	1	2	3	4	5
Error	0.329	1.468	2.942	8.878	7.421
Variance	0.053	10.894	13.834	112.268	40.258
Uniform noise	1	2	3	4	5
Error	0.128	0.116	0.152	0.207	0.266
Variance	0.105	0.007	0.007	0.017	0.028

Table 1: The results of experiments with simulated data.

where $\mathcal{A}_3$ is the group of affine transformations on $\mathcal{P}^3$. In words: we compute the affine transformation which maps the affine reconstruction as closest possible to the Euclidean one (i.e. which minimizes the sum of absolute point distances). The resulting sum of absolute point distances (in Euclidean space), divided by the number of reconstructed points, serves as quality value of the affine reconstruction.

7.3 Experiments with simulated data

We effectuated the simulations for 9 virtual cameras in different positions and considered each of the camera pairs as an own stereo system. So, we obtained 36 different virtual stereo systems. For each of them we effectuated the following experiments:

1. A plane with 92 points of interest is translated 4 times in the same direction. Those 5 sets of each 92 coplanar points are projected by the two cameras.
2. Perturbation of the image points by gaussian or uniform noise from 1 to 5 pixels variance.
3. Projective calibration, using the perturbed image points.
4. Affine calibration, using the perturbed image points.
5. Projection of a set of 60 object points.
6. Perturbation of the image points by gaussian or uniform noise from 1 to 5 pixels.
7. Affine reconstruction of the 60 points using the affine calibration obtained in step 4.
8. Evaluation of the affine reconstruction using the criterion defined in paragraph 7.2.

Table 1 shows the obtained results where each of the error values is the mean of 36 different experiments with 60 object points. The error is given in affine unit which can be related to the size of the min-max-cube of the object points, $30 \times 41 \times 22.5$.

7.4 Experiments with real data

The sequence of the experiments' steps is the one described in paragraph 7.1. The image points used for affine calibration have been obtained by taking 5 pairs of images of a calibration grid (see Fig 5), which was subject to a translational movement between each taking, but always in the same direction. It is important to note, that the parallelism of the line segments in the calibration grid was not considered for affine calibration. Only the pure information of coplanarity has been used. The images of the calibration grid have been independently used to establish a Euclidean calibration of the stereo system. As scenes to reconstruct we used objects like those in Table 2.

We have effectuated the affine and the Euclidean calibrations for 8 different positions of the calibration grid. The values in Table 2 show the errors of the affine reconstructions in comparison to the Euclidean ones (see paragraph 7.2). Each value is the mean for the 8 different calibration setups. The error values are given in *cm*. In order to well interpret the error values it is important to note, that the height of the house, which has been observed in scenes 2, 3, and 4, is about 30 cm.

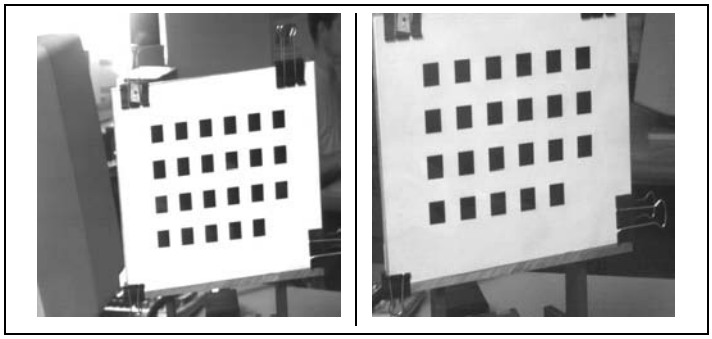


Figure 5: Two images of the calibration grid which was used in order to determine image correspondences of coplanar points.

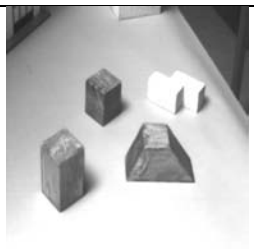
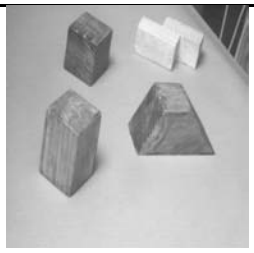
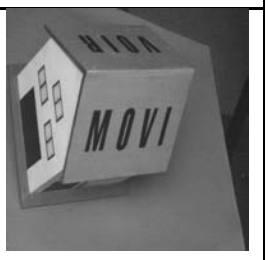
	Scene 1	Scene 2	Scene 3	Scene 4
Error [cm]	0.072	0.213	0.173	0.274
Variance [cm]	0.001	0.032	0.017	0.027
Image 1				
Image 2				

Table 2: The results of experiments with real images. Errors and variances are given in cm.

Another important remark is, that the images have been taken at distances inferior to 1.5 m. So, the following condition holds:

$$\frac{\text{distance object-camera}}{\text{object's size}} < 5$$

Normally, under this condition, the assumption of an affine camera model would give bad results. Since we are using an entirely projective model, our affine calibration method has no trouble with this condition.

In Fig 6 two examples of affine reconstructions of scenes 1 and 2 are shown. Corresponding Euclidean reconstructions are shown in order to illustrate the "affine deformation" characterizing the affine reconstructions.

8 Simulation of a vehicle equiped with affinely calibrated cameras

We present a simple example for a possible application of affine calibration. Fig 7 shows a road that has to be pursued by a vehicle. The simple strategy of the vehicle is to always try to keep the middle of the road. Since "middle" or mid-points are affine properties, an affine calibration of a stereo system mounted on the vehicle should be enough to perform the task.

The vehicle is represented in Fig 8 as a bold point, and a line which shows its orientation. The borders of the

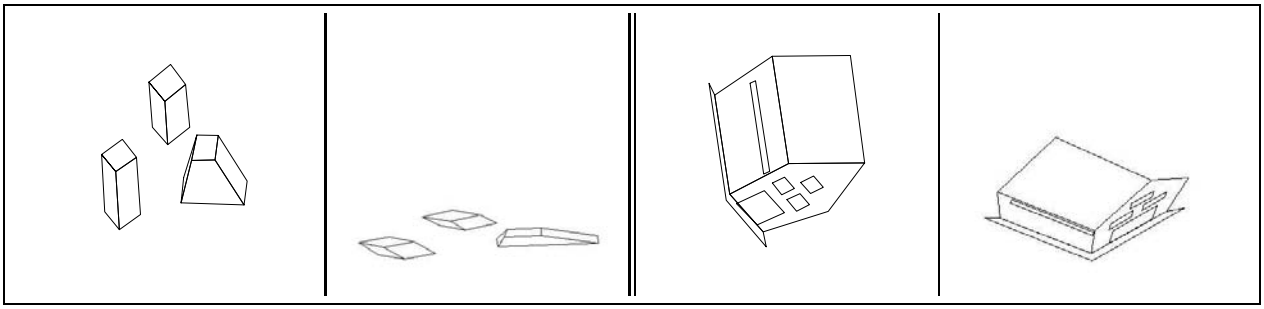


Figure 6: Euclidean and affine reconstructions of two scenes.

road are formed by sets of feature points which are observed by the two cameras. In the simulation, the cameras observe at each instant only 2 points of each border. In addition, the cameras also capture two reference points of the vehicle which indicate its orientation.

At each instant the middle of the road can be determined in the images. For doing so, the mid-points of two pairs of facing border points are detected, using the infinity homography and simple calculations (see for instance [11]). In each image, the middle of the road is then represented as a straight line. By comparing the relative position of the vehicle's reference points with the line representing the middle of the road, commands to the vehicle's guiding system can be created in order to keep the vehicle in the middle.

Some steps of the simulated trip are shown in Fig 8. In part (b) we see that the car weakly drifts away from the middle. This is because the cameras don't see very "far", only two border points ahead, and so the vehicle reacts a little too late in order to well follow the curve.

In part (c) it is shown that the vehicle re-gained the middle of the road. In addition to the primary task of the vehicle to pursue the road, a secondary task of obstacle detection is implemented. Here, the road is taken as reference plane for a quasi-affine calibration of the camera system (see paragraph 3.2). This means nothing else that, using the border points, the plane collineation for the ground plane is determined. This plane collineation enables detection of points which are above the ground plane, and which represent obstacles in the actual context.

In Fig 8 (c) an obstacle is detected. The simple obstacle avoiding strategy applicated here is, that the vehicle slows down until it stops, and that then the obstacle is removed (Fig 8 (d)).

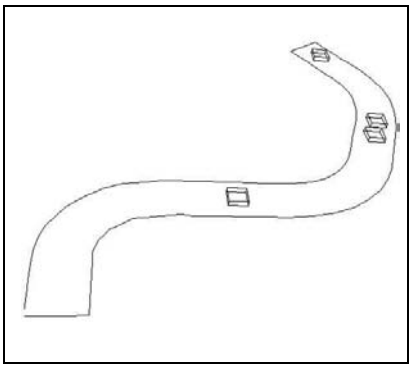


Figure 7: 3D representation of the road to pursue. 4 obstacles are visible.

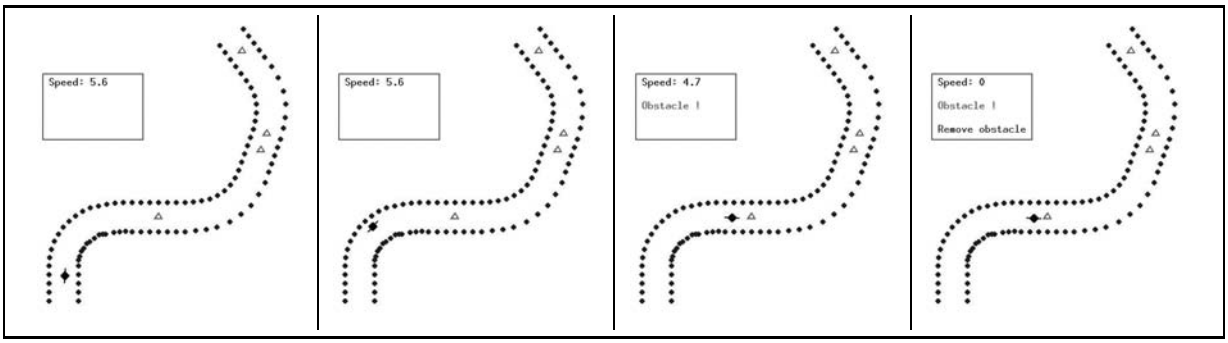


Figure 8:
(a) The vehicle keeps the middle of the road. **(b)** In the curve the vehicle weakly drifts away from the middle. **(c)** An obstacle is detected; the vehicle slows down. **(d)** The vehicle stops; the obstacle will be removed.

9 Discussion

In this paper, we have firstly reviewed some concepts of partial stereo calibration and affine reconstruction. We have examined affine calibration for cameras with same intrinsic parameters and have shown that in this case only 2 instead of 3 correspondences of vanishing points are needed.

We have proposed two practical methods of general affine stereo calibration which can directly be applied for affine self-calibration. Only for the calibration step a translational movement of the camera system would be necessary.

Our approach goes in the same direction as the one proposed in [1] where the camera system is supposed to move only in a single plane and to rotate around axes which are perpendicular to this plane. With that approach, which is also well applicable for vehicles, 3 stereo pairs are needed for affine calibration whereas our approach needs only 2. This is due to our assumption of a translational movement which is a little bit stronger than the one used in [1].

We have experimented one of the proposed affine calibration methods with simulated as well as real data and the results show the applicability of the approach. A simulation of a vehicle which is equiped with an affinely calibrated stereo system, illustrates one possible application of the concept of affine calibration.

References

[1] Beardsley, P.A., Zisserman, A.: Affine Calibration of Mobile Vehicles. Proc. of the Workshop on Geometrical modelling and invariant computer vision, China, 1995, to appear

[2] Faugeras, O.: What can be seen in three dimensions with an uncalibrated stereo rig?. Proc. of the 2nd ECCV, Santa Margherita Ligure, Italy, 1992, 563–578

[3] Hartley, R., Sturm, P.: Triangulation. Proc. of the ARPA Image Understanding Workshop 1994, Monterey, CA, 1994

[4] Hartley, R.I.: Euclidean reconstruction from uncalibrated views. Proc. of the Darpa–Esprit workshop on Applications of Invariants in Computer Vision, Azores, Portugal, 1993, 187–202

[5] Hartley, R.: Invariants of points seen in multiple images. Technical Report, G.E. CRD, Schenectady, 1992

[6] Hartley, R., Gupta, R., Chang, T.: Stereo from uncalibrated cameras. Proc. of the Conf. on Computer Vision and Pattern Recognition, Urbana-Champaign, Illinois, USA, 1992, 761–764

[7] Koenderink, J.J., van Doorn, A.J.: Affine structure from motion. Technical Report, Utrecht University, The Netherlands, 1989

[8] Luong, Q.T., Vieville, T.: Canonic Representations for the Geometries of Multiple Projective Views. Technical Report, University of California, Berkeley, 1993

[9] Mohr, R., Quan, L., Veillon, F.: Relative 3D reconstruction using multiple uncalibrated images. The International Journal of Robotics Research, 1995, to appear

[10] Quan, L.: Affine stereo calibration for relative affine shape reconstruction. Proc. of the 4th British Machine Vision Conference, Surrey, England 1993, 659–668

[11] Quan, L., Mohr, R.: Affine shape representation from motion through reference points. Journal of Mathematical Imaging and Vision (1), 1992, 145–151

[12] Semple, J.G., Kneebone, T.: Algebraic Projective Geometry. Oxford Science Publication, 1952

[13] Shashua, A.: Projective Structure from Two Uncalibrated Images : Structure from Motion and Recognition. Technical Report A.I. Memo No. 1363, MIT, 1992

A Detection of vanishing points and lines

Vanishing points and lines play an important role in affine calibration and reconstruction. We describe here several methods to detect vanishing points or lines.

A.1 Parallel planes and lines

Parallel planes and lines intersect in the infinity. Thus, the projections of intersection points respectively lines are vanishing points respectively lines. Based on a priori knowledge about parallelism of observed planes or lines, we can determine these vanishing points or lines. We indicate how to proceed, depending on what is observed.

Parallel lines Determination of the vanishing point consists on determination of the point which best approximates the intersection of the projected lines (see paragraph 4.1).

Parallel planes See the methods described in paragraph 4.2 and section B.

Plane and line which are parallel Let H be the plane collineation induced by the considered plane and $\mathbf{d}$ and $\mathbf{d}'$ the two projections of the line. Then, the vanishing points are [10]

$$\begin{aligned} \mathbf{q}' &\approx \mathbf{d}' \wedge (H^* \mathbf{d}) \\ \mathbf{q} &\approx H^{-1}\mathbf{q}' \end{aligned}$$

where H^* is the dual transformation of H.

A.2 Ratio and double ratio

Consider three collinear 3D points whose ratio is known. Taking the point at infinity of the line formed by these points as $4th$ point, their double ratio equals the known ratio (up to an ambiguity in the order of points in double ratio computation). Since the double ratio of 4 collinear points is conserved by any projective transformation, the projection of the point at infinity can be determined, being given the projections of the 3 collinear points and their ratio in 3D.

A.3 Parallel conics

See the descriptions in sections C and D.

B Intersection of planes

Consider two planes in 3D space whose plane homographies between two image planes are known. We describe here, how the projections of the intersection line of the planes in the two images can be determined from the plane homographies. We show as well that the two epipoles can be determined in non-degenerated cases. The method described here stands in contrast to the one described in paragraph 4.2. But, we describe it here only for theoretic interests.

In the following, we suppose that neither of the two planes contains any of the centers of projections of the two cameras.

Let

- Π_1 and Π_2 the two planes
- H_1 and H_2 the homographies induced by the two planes
- D the intersection line of Π_1 and Π_2
- $\mathbf{d}$ and $\mathbf{d}'$ the projections of D.

Consider any point Q on the intersection line D. For its projections, $\mathbf{q}$ and $\mathbf{q}'$, we have

$$H_1\mathbf{q} \approx \mathbf{q}' \approx H_2\mathbf{q}.$$

Thus, there is a $\lambda \in \mathcal{R}$ with

$$(H_1 - \lambda H_2)\,\mathbf{q} = 0. \tag{8}$$

This equation is valid when taking as $\mathbf{q}$ any of the following but no other points:

- the projection of any point of D
- the epipole in the first image, $\mathbf{e}$

For that equation 8 has non-trivial solutions for λ we must have

$$det\,(H_1 - \lambda H_2) = 0. \tag{9}$$

This is a polynomial equation of $3rd$ degree in λ. Its 3 solutions are λ_1, λ_2, and λ_3 (the generalized eigenvalues of the matrix family defined by H_1 and H_2).

Since there is one line, whose points all fulfill equation 8 (the line d'), one of the ratios of equation 9 must be double. Thus, let $\lambda_1 = \lambda_2$. It follows that λ_1 and thus also λ_2 are real numbers. Consequently, λ_3 is a real number[6].

Concluding the several remarks, we obtain:

1. The line $\mathbf{d}$ is the kernel of the matrix $(H_1 - \lambda_1 H_2)$.
2. We obtain $\mathbf{d}'$ by $\mathbf{d}' \approx H_1^*\mathbf{d} \approx H_2^*\mathbf{d}$.
3. The first epipole, $\mathbf{e}$, is the kernel of $(H_1 - \lambda_3 H_2)$.
4. The second epipole, $\mathbf{e}'$, is determined by $\mathbf{e}' \approx H_1\mathbf{e} \approx H_2\mathbf{e}$.

These conclusions are all valid only in the non-degenerated case, when neither $\mathbf{d}$ contains $\mathbf{e}$ nor $\mathbf{d}'$ contains $\mathbf{e}'$. In the following paragraph we make the distinction between non-degenerated and degenerated cases.

B.1 Different cases

Sample graphs of the generalized characteristical polynomial of $H_1 - \lambda H_2$ are shown in Fig 9 (a) and (b) for the two following cases.

Case 1 : $\lambda_1 \neq \lambda_3$. That means that neither $\mathbf{e}$ is on $\mathbf{d}$ nor $\mathbf{e}'$ on $\mathbf{d}'$. This is the general case, in which $\mathbf{d}$ and $\mathbf{d}'$ as well as $\mathbf{e}$ and $\mathbf{e}'$ can be determined.

Case 2 : $\lambda_1 = \lambda_3$. This is the degenerated case in which we meet the situation that $\mathbf{e}$ lies on $\mathbf{d}$ and $\mathbf{e}'$ on $\mathbf{d}'$. We are able to determine $\mathbf{d}$ and $\mathbf{d}'$ but we can not identify $\mathbf{e}$ and $\mathbf{e}'$ on these lines.

B.2 Projective calibration

Suppose that we are in the non-degenerated case 1. So, we can determine the epipoles, and also the Fundamental Matrix F:

$$F \approx [\mathbf{e}']_\times H_1 \approx [\mathbf{e}']_\times H_2$$

[6] A polynomial has always an even number (or 0) of imaginary ratios

B.3 The generalized eigenvalues in presence of noise

This method for the determination of the projections of the intersection of two planes is of course sensitive to noise. In the following we clarify what we mean by noise in the present context.

Consider a system of two cameras. Plane homographies can not have any form [8], especially they always must map the first epipole onto the second one. We thus talk of noise, when we have plane homographies H_1 and H_2 which are "not compatible" with the stereo system. This indeed occurs in practice, where the "exact" homographies can only be approximated.

In the presence of noise, the generalized characteristical polynomial in equation 9 does in general not have a double ratio. Thus, $\mathbf{d}$, $\mathbf{d}'$, $\mathbf{e}$, and $\mathbf{e}'$ can not be determined uniquely. Sample graphs of the generalized characteristical polynomial of $H_1 - \lambda H_2$ in the presence of noise are shown in Fig 9 (c) and (d).

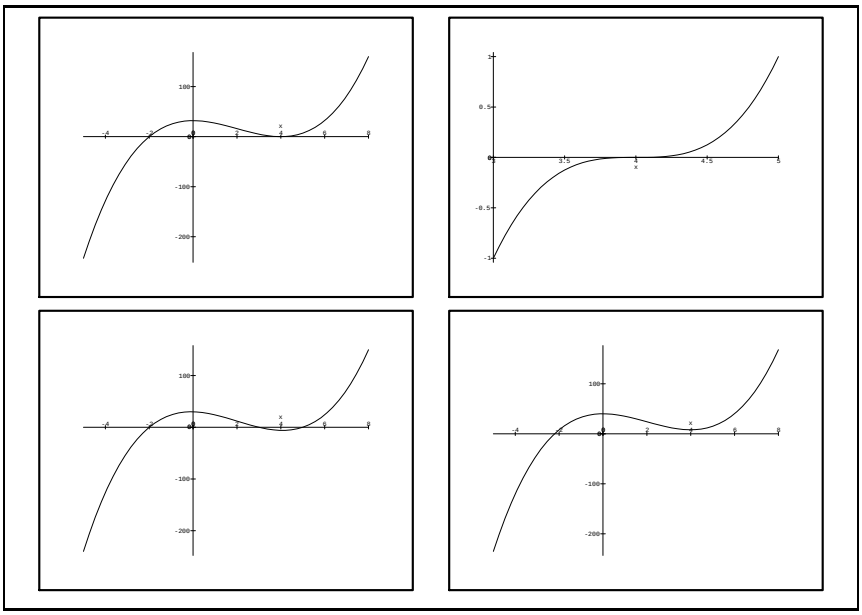


Figure 9: These graphs illustrate several cases for the graph of the generalized characteristical polynomial of $H_1 - \lambda H_2$.
(a) general case without noise: there are two real ratios, one is a double one. This means that the epipoles do not lie on the projections of the intersection line D of the planes. **(b)** (on the top at right) degenerated case without noise: there is one triple real ratio. The epipoles lie on the respective projection of D. **(c)** with noise: there are three distinct real ratios. **(d)** with noise: there are one real and two imaginary ratios. The imaginary ratios are of course complex conjugates.

C Parallel 3D circles

We describe how vanishing lines can be determined, when "parallel" 3D circles are observed. By parallel circles we mean circles whose supporting planes are parallel.

C.1 Reflections for a single camera

Remember that every two circles in the affine plane intersect in the two co-circular points I and J (see paragraph 2.5). Since I and J are points at infinity, it follows that parallel planes have common co-circular points. Thus, all parallel circles intersect in the two co-circular points which are common to the supporting planes.

In the following we restrain us to the case of two parallel circles. The only condition we pose is that the projetions of the circles are not identical.

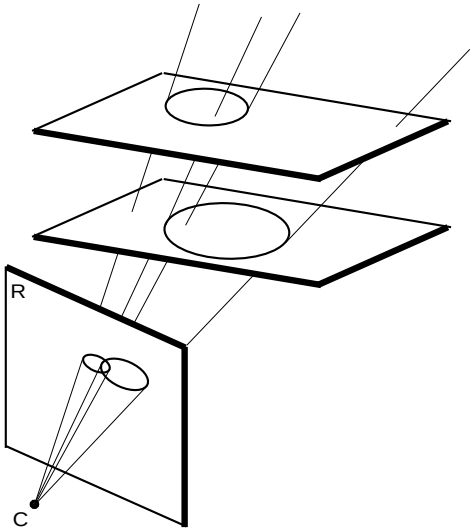


Figure 10: Perspective projection of two parallel circles

Consider now a perspective projection, which is represented by a matrix $P_{3\times 4}$. The circles are projected on conic sections $\mathcal{C}_\rangle$ (see Fig 10), with matrix representations C_i. Let in the following be $\mathbf{q}_i, i = 1, \ldots, 4$ the 4 intersection points of $\mathcal{C}_\infty$ and $\mathcal{C}_\in$.

We now want to determine the projections of the two co-circular points I and J which are common to the supporting planes of the 3D circles. Let $\mathbf{c}_1$ and $\mathbf{c}_2$ be these projections. Concerning the composition of the 4 intersections points $\mathbf{q}_i$, we can distinguish 3 cases:

1. The real parts of $\mathcal{C}_\infty$ and $\mathcal{C}_\in$ intersect in two distinct points, say $\mathbf{q}_1$ and $\mathbf{q}_2$.

   In this case there are also two imaginary intersection points of $\mathcal{C}_\infty$ and $\mathcal{C}_\in$, say $\mathbf{q}_3$ and $\mathbf{q}_4$. Since any perspective projection maps imaginary points on also imaginary points, we have: $\{\mathbf{c}_1, \mathbf{c}_2\} = \{\mathbf{q}_3, \mathbf{q}_4\}$, i.e. the projections of the co-circular points are the two imaginary intersection points of the conics.

2. The real parts of $\mathcal{C}_\infty$ and $\mathcal{C}_\in$ are tangent in one point.

   So, we have one double real intersection $\mathbf{q}_1 = \mathbf{q}_2$ and two imaginary ones, $\mathbf{q}_3$ and $\mathbf{q}_4$. By the same argument as in case 1 we conclude that $\{\mathbf{c}_1, \mathbf{c}_2\} = \{\mathbf{q}_3, \mathbf{q}_4\}$.

3. The real parts of $\mathcal{C}_\infty$ and $\mathcal{C}_\in$ do not intersect.

   The 4 intersection points are thus imaginary. Furthermore, there are two pairs of complex conjugated points (see paragraph C.2), say $(\mathbf{q}_1, \mathbf{q}_2)$ and $(\mathbf{q}_3, \mathbf{q}_4)$. Since any perspective projection maps complex conjugated points on also complex conjugated points (see paragraph C.2), we know $\mathbf{c}_1$ and $\mathbf{c}_2$ up to one ambiguity. We have indeed the two possibilities $\{\mathbf{c}_1, \mathbf{c}_2\} = \{\mathbf{q}_1, \mathbf{q}_2\}$ and $\{\mathbf{c}_1, \mathbf{c}_2\} = \{\mathbf{q}_3, \mathbf{q}_4\}$. In C.6 we show how this ambiguity can be solved for.

C.2 Proofs

Intersection of two conics.

It is known that two conic sections intersect in 4 points which can be imaginary and/or multiple [12].

We show here by a simple argument that, if the intersection of two conics contains an imaginary point then it also contains its complex conjugated point.

Proof.

A point $\mathbf{q}$ lies on the conic $\mathcal{C}$, if

$$\mathbf{q}^T C \mathbf{q} = 0 \quad .$$

This is a polynomial of $2nd$ degree in the coefficients of $\mathbf{q}$. Thus, the equation rests true if we put for every coefficient of $\mathbf{q}$ its complex conjugated number. It follows that the complex conjugated point of $\mathbf{q}$ lies on $\mathcal{C}$. QED.

Linear transformation of complex conjugated points

We show that linear transformations map complex conjugated points on also complex conjugated points.

Proof. Let $Q, \overline{Q} \in \mathcal{C}^n$ complex conjugated points and $M_{m\times n}$ a real matrix[7].

We have for all j, j=1,...,m:

$$(MQ)_j = \sum_{k=1}^{n} M_{jk}Q_k = \overline{\sum_{k=1}^{n} M_{jk}\overline{Q_k}} = \overline{(M\overline{Q})_j}.$$

Thus

$$MQ = \overline{M\overline{Q}}.$$

QED.

C.3 Reflections for two cameras

Suppose that we do the just described observations for two cameras. We so dispose for each camera of the projections of the two co-circular points which are common to the planes supporting the 3D circles. Let $\mathbf{c}_1$ and $\mathbf{c}_2$ be the projections in the first image and $\mathbf{c}'_1$ and $\mathbf{c}'_2$ those in the second image.

What we do not yet know is the correspondence between these points, i.e. we do not know which of $\mathbf{c}'_1$ and $\mathbf{c}_2$ are the corresponding images to $\mathbf{c}_1$ and $\mathbf{c}_2$. We use the epipolar constraint in order to obtain this correspondence. Thus, the corresponding point to $\mathbf{c}_1$ is the one of $\mathbf{c}'_1$ and $\mathbf{c}'_2$ which lies on the epipolar line $F\mathbf{c}_1$.

We suppose here that $\mathbf{c}'_1$ and $\mathbf{c}'_2$ do not lie on the same epipolar line. This can only occur if the intersection line of the supporting planes lies completely in an epipolar plane.

In the following we suppose that $(\mathbf{c}_1, \mathbf{c}'_1)$ and $(\mathbf{c}_2, \mathbf{c}'_2)$ are the corresponding images of the co-circular points.

C.4 Affine calibration - Determination of the infinity homography

In the previous sections we have shown how the projections of co-circular points of parallel planes can be determined. Since co-circular points are points at infinity, their projections are consequently vanishing points. Thus, they can be used for determination of the infinity homography.

Consider the two pairs of corresponding projections of co-circular points $(\mathbf{c}_1, \mathbf{c}'_1)$ and $(\mathbf{c}_2, \mathbf{c}'_2)$. Only one of the pairs contributes to the determination of the infinity homography, since $\mathbf{c}_1$ and $\mathbf{c}_2$ respectively $\mathbf{c}'_1$ and $\mathbf{c}'_2$ are complex conjugates.

But, we already dispose of the correspondence between the two vanishing lines $\langle \mathbf{c}_1, \mathbf{c}_2 \rangle$ and $\langle \mathbf{c}'_1, \mathbf{c}'_2 \rangle$. Together with the epipolar geometry we can thus obtain further correspondences of vanishing points. To do so, choose one point on $\langle \mathbf{c}_1, \mathbf{c}_2 \rangle$. Its corresponding vanishing point is just the intersection of $\langle \mathbf{c}'_1, \mathbf{c}'_2 \rangle$ and the epipolar line in the second image.

So, by observation of parallel 3D circles, we are able to obtain two correspondences of vanishing points which fully contribute to the determination of the infinity homography. Thus, two non-parallel pairs of parallel circles are sufficient to determine the infinity homography or, equivalently, to do affine calibration.

C.5 Euclidean calibration

The Euclidean calibration of a camera is equivalent to the determination of the projection of the absolute conic (see for example [8]). The projection of the absolute conic is itself a conic and thus defined uniquely by 5 points in general position. Since the absolute conic is formed just by the co-circular points of all planes of $\mathcal{P}^3$, observation of 3 non-parallel pairs of parallel circles is sufficient for determination of the absolute conic's projection, and thus for Euclidean calibration.

C.6 Resolving for the ambiguity in case 3

We briefly describe two possible methods which permit to resolve for the ambiguity in case 3 (paragraph C.1).

[7] Each linear transformation can be represented by such a matrix.

Three or more circles. When observing more than two parallel circles, we choose the pair of complex conjugate points which appears in every set of intersection points of respectively two conics.

Two or more images. Consider a second projection of the two circles (see Fig 11). If also in this second image the real parts of the projected conics do not have any any common point, we are still not able to resolve for the ambiguity. But, if we meet the contrary case, we can determine the projections of the co-circular points in the second image. Using the epipolar geometry, we then can identify the projections of the co-circular points in the first image.

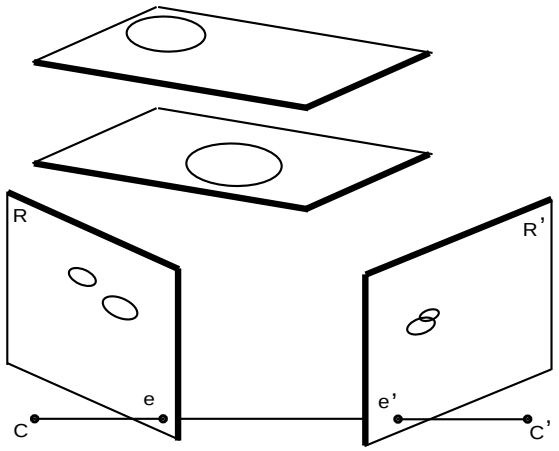


Figure 11: In the left image, the conics do not have any real intersection point, in the right image there are two of them.

D Parallel 3D parabolas

We describe how vanishing lines can be determined, when "parallel" 3D parabolas are observed.

Two conic sections in the affine plane have two common tangents, one of which is the line at infinity of the affine plane. The projection of a tangent line on a 3D conic is tangent to the projected conic. Thus, by observing two coplanar 3D parabolas, we can determine the vanishing line of their supporting plane up to one ambiguity. Of course this is also valid for two parallel 3D parabolas. When we observe more than two coplanar parabolas in general position, the vanishing line of their supporting planes can be uniquely determined. It is just the single line which is tangent to all of the projected conics.

Once corresponding vanishing lines are determined for two cameras, we obtain two correspondences of vanishing points like described in paragraph C.4. We conclude that the observation of two non-parallel triplets of parallel parabolas is sufficient for the determination of the infinity homography.