

# Closed-form Solutions for the Euclidean Calibration of a Stereo Rig

Gabriela Csurka, David Demirdjian, Andreas Ruf, Radu Horaud

► **To cite this version:**

Gabriela Csurka, David Demirdjian, Andreas Ruf, Radu Horaud. Closed-form Solutions for the Euclidean Calibration of a Stereo Rig. 5th European Conference on Computer Vision (ECCV '98), Jun 1998, Freiburg, Germany. pp.426–442, 10.1007/BFb0055682 . inria-00590090

**HAL Id: inria-00590090**

**<https://hal.inria.fr/inria-00590090>**

Submitted on 3 May 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Closed-form solutions for the Euclidean calibration of a stereo rig

G. Csurka, D. Demirdjian, A. Ruf, and R. Horaud

INRIA Rhône-Alpes, 655 Av. de l'Europe, 38330 Montbonnot Saint Martin, France

**Abstract.** In this paper we describe a method for estimating the internal parameters of the left and right cameras associated with a stereo image pair. The stereo pair has known epipolar geometry and therefore 3-D projective reconstruction of pairs of matched image points is available. The stereo pair is allowed to move and hence there is a collineation relating the two projective reconstructions computed before and after the motion. We show that this collineation has similar but different parameterizations for general and ground-plane rigid motions and we make explicit the relationship between the internal camera parameters and such a collineation. We devise a practical method for recovering four camera parameters from a single general motion or three camera parameters from a single ground-plane motion. Numerous experiments with simulated, calibrated and natural data validate the calibration method.

## 1 Introduction

Traditional stereo vision systems use a single image pair to provide projective, affine, or Euclidean reconstruction. It has been clear that redundancy offered by further image pairs can significantly increase the quality and stability of visual reconstructions. Nevertheless, if the visual task is to recover metric structure, there are problems because both the intrinsic parameters (of the left and right cameras) and extrinsic ones (relative position and orientation of the two cameras) can vary over time. This is particularly critical when an active stereo head is being used. It is therefore important to be able to re-calibrate the stereo rig over time and over a small number of motions without using any special purpose calibration device.

A number of authors investigated the relationship between projective, affine, and metric spaces in conjunction with a single camera undergoing rigid motions [10,6,7,13,15]. In [14] it is shown that there are many critical situations for which metric structure cannot be recovered. When image pairs are available one may use additional constraints. Affine structure can be recovered from either pure translations [12], pure rotations [2] or ground-plane motions [1] of a stereo rig. For general motions, affine structure can be estimated from the eigenvector of a 3-D collineation [17] and metric structure can be estimated from two general motions [3]. Furthermore, in [8] it is shown that affine structure is an intrinsic property of a rigid stereo rig and that it can be easily recovered by combining any number of general motions.

In this paper we develop closed-form solutions for the self-calibration of a stereo rig from a **single** general or ground-plane motion. The basic assumption is that the stereo rig has the same internal and external parameters before and after the motion. More precisely, let  $\mathcal{P}_1$  and  $\mathcal{P}_2$  be two projective reconstructions obtained with an uncalibrated stereo rig before and after a rigid motion. These two reconstructions, i.e., a set of 3-D points, are related by a  $4 \times 4$  collineation  $\mathbf{H}_{12}$  which is related to the rigid motion  $\mathbf{D}_{12}$  by ([17,3]):

$$\mathbf{H}_{12} \simeq \mathbf{H}_{PE}^{-1} \mathbf{D}_{12} \mathbf{H}_{PE} \quad (1)$$

As it will be shown below, an immediate consequence of this similarity relationship is that  $\mathbf{H}_{12}$  can be factorized as  $\mathbf{A}\mathbf{J}\mathbf{A}^{-1}$  where  $\mathbf{J}$  is a real Jordan canonical form. We prove that this factorization is not unique and we show how to parameterize all such factorizations and how to estimate them in practice. Furthermore, we show that there exists a relationship between the left (or right) camera parameters and all possible factorizations of  $\mathbf{H}_{12}$ .

More precisely, let  $\mathbf{K}$  be the upper triangular matrix associated with the left camera such that  $K_{33} = 1$ . We will show that (i) in the case of a general motion the matrix  $\mathbf{K}\mathbf{K}^\top$  is parameterized by a pencil of conics and that (ii) for a ground-plane motion  $\mathbf{K}\mathbf{K}^\top$  is parameterized by a linear combination of three conics. Therefore, one constraint onto the entries of  $\mathbf{K}$  is sufficient to estimate four intrinsic parameters from a single motion and two constraints onto the entries of  $\mathbf{K}$  are necessary for estimating three intrinsic parameters from a single planar motion. As a consequence, a camera with zero image skew can be calibrated from a general motion and a camera with zero image skew and known aspect ratio can in turn be calibrated from one ground-plane motion.

The method described in this paper has several advantages over previous stereo calibration approaches. The first advantage is that affine calibration is not necessary prior to metric calibration as it is done with stratified approaches. This is particularly important for ground-plane motions for which affine structure has proved difficult to obtain. The second advantage is that all the computations are based on linear algebra techniques such as singular value decomposition. The third advantage is that, while a single motion is sufficient to calibrate, several motions can be combined in conjunction with a standard outliers rejection method in order to estimate the calibration parameters more robustly.

## 1.1 Paper organization

The remainder of the paper is organized as follows. Section 2 briefly recalls the camera model and makes explicit the structure of  $\mathbf{K}\mathbf{K}^\top$  for a camera with zero skew and for a camera with zero skew and known aspect ratio. Section 3 recalls the mathematical properties associated with equation (1). Section 4 describes the real Jordan factorization of matrix  $\mathbf{H}_{12}$ , analyses this factorization from a geometrical point of view, shows how to parameterize all possible factorizations, and describes a method to compute these factorizations in practice. Section 5 shows how to perform Euclidean calibration from the real Jordan factorization

of  $\mathbf{H}_{12}$  for general and ground-plane motions. Section 6 validates the method with simulated and real data and Section 7 discusses the method in the light of the experimental results obtained so far.

## 2 Camera model and the absolute conic

A pinhole camera projects a point  $\mathbf{M}$  from the 3-D projective space onto a point  $\mathbf{m}$  of the 2-D projective plane. This projection can be written as  $\mathbf{m} \simeq \mathbf{P}\mathbf{M}$ , where  $\mathbf{P}$  is a  $3 \times 4$  homogeneous matrix of rank equal to 3 and the sign  $\simeq$  designates the projective equality – equality up to a scale factor. If we restrict the 3-D projective space to the Euclidean space, then it is well known that  $\mathbf{P}$  can be written as:

$$\mathbf{P}_E = \mathbf{K} (\mathbf{R} \ t) = (\mathbf{K}\mathbf{R} \ \mathbf{K}t) \quad (2)$$

where  $\mathbf{R}$  and  $t$  describe the orientation and the position of the camera in the chosen Euclidean frame. If we consider the standard camera frame as the 3-D Euclidean frame (the origin is the center of projection, the xy-plane is parallel to the image plane and the z-axis points towards the visible scene), the projection matrix becomes<sup>1</sup>  $\mathbf{P}_E = (\mathbf{K} \ \mathbf{0}_3)$ .

The most general form for the matrix of intrinsic parameters  $\mathbf{K}$  is:

$$\mathbf{K} = \begin{pmatrix} \alpha & r & u_0 \\ 0 & k\alpha & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (3)$$

where  $\alpha$  is the horizontal scale factor,  $k$  is the ratio between the vertical and horizontal scale factors,  $r$  is the image skew and  $u_0$  and  $v_0$  are the image coordinates of the center of projection.

The relation between the matrix  $\mathbf{K}$  and the image of the absolute conic is  $\mathbf{C} \simeq \mathbf{K}^{-\top} \mathbf{K}^{-1}$  [4]. Let us make explicit the dual of this conic, i.e.,  $\mathbf{A} = \mathbf{C}^{-\top}$ :

$$\mathbf{A} \simeq \mathbf{K}\mathbf{K}^\top = \begin{pmatrix} \alpha^2 + r^2 + u_0^2 & rk\alpha + u_0v_0 & u_0 \\ rk\alpha + u_0v_0 & k^2\alpha^2 + v_0^2 & v_0 \\ u_0 & v_0 & 1 \end{pmatrix} \quad (4)$$

This means that  $\mathbf{A}$  is symmetric and if we want to fix the scale factor ( $\mathbf{A}$  is defined up to a scale) such that  $\mathbf{A} = \mathbf{K}\mathbf{K}^\top$ , we need that  $A_{33} = 1$ .

Equation (3) describes a five-parameter camera. It will be useful to consider camera models with a reduced set of intrinsic parameters, as follows:

- *four-parameter camera* with  $r = 0$  (image skew), which means that the image is a rectangle – a sensible assumption. In this case the dual conic becomes:

$$\mathbf{A} \simeq \mathbf{K}\mathbf{K}^\top = \begin{pmatrix} \alpha^2 + u_0^2 & u_0v_0 & u_0 \\ u_0v_0 & k^2\alpha^2 + v_0^2 & v_0 \\ u_0 & v_0 & 1 \end{pmatrix} \quad (5)$$

---

<sup>1</sup> We denote by  $\mathbf{0}_n$  the  $n$ -vector containing  $n$  zeros.

which provides an additional constraint on the entries of  $\mathbf{A}$ , i.e.  $A_{12} - A_{13}A_{23} = 0$ .

- *three-parameter camera* with  $r = 0$  and  $k$  (aspect ratio) having a known value; for instance the value of  $k$  can be obtained from the physical size of a pixel. Therefore there is an additional constraint on the entries of  $\mathbf{A}$ :

$$k^2(A_{11} - A_{12}^2) - (A_{23}^2 - A_{22}) = 0$$

### 3 Rigid motion of a stereo rig

A stereo rig is composed of two cameras fixed together. Let  $\mathbf{P}$  and  $\mathbf{P}'$  be the projection matrices of the left and right cameras. We can choose without loss of generality a projective basis  $\mathcal{B}_P$  such that<sup>2</sup>  $\mathbf{P} = (\mathbf{I}_3 \mathbf{0}_3)$ . In this case  $\mathbf{P}' = (\mathbf{H}_\infty + \mathbf{e}'\mathbf{a}^\top, \lambda\mathbf{e}')$  [11], where  $\mathbf{H}_\infty$  is the infinite homography between the left and right images,  $\mathbf{e}'$  is the right epipole,  $\mathbf{a}$  an arbitrary 3-vector and  $\lambda$  is an arbitrary scale factor. It was shown in [8] that the 4-vector  $(\mathbf{a}^\top \lambda)$  has a simple but important geometric interpretation, namely it is the plane of infinity associated with the stereo pair. However, this plane is not used throughout this paper.

Given a stereo rig with two projection matrices  $\mathbf{P}$  and  $\mathbf{P}'$ , it is possible to compute the 3-D projective coordinates of a point  $\mathbf{M}$  in the basis  $\mathcal{B}_P$  from the equations  $\mathbf{m} \simeq \mathbf{P}\mathbf{M}$  and  $\mathbf{m}' \simeq \mathbf{P}'\mathbf{M}$ , where  $\mathbf{m}$  and  $\mathbf{m}'$  are the projections of  $\mathbf{M}$  onto the left and right images.

If we restrict the projective space to the Euclidean space and choose as basis  $\mathcal{B}_E$  the standard camera frame associated with the first camera,  $\mathbf{P}$  and  $\mathbf{P}'$  are given by:

$$\mathbf{P}_E = (\mathbf{K} \mathbf{0}_3) \quad \text{and} \quad \mathbf{P}'_E = (\mathbf{K}'\mathbf{R} \mathbf{K}'\mathbf{t})$$

where  $\mathbf{K}$  and  $\mathbf{K}'$  are the matrices of left and right intrinsic camera parameters,  $\mathbf{R}$  and  $\mathbf{t}$  describe the orientation and position of the right camera frame with respect to the left camera frame. The equations  $\mathbf{m} \simeq \mathbf{P}_E\mathbf{M}_E$  and  $\mathbf{m}' \simeq \mathbf{P}'_E\mathbf{M}_E$  allow the estimation of  $\mathbf{M}_E$  – the 3-D Euclidean coordinates of a 3-D point in the basis  $\mathcal{B}_E$ .

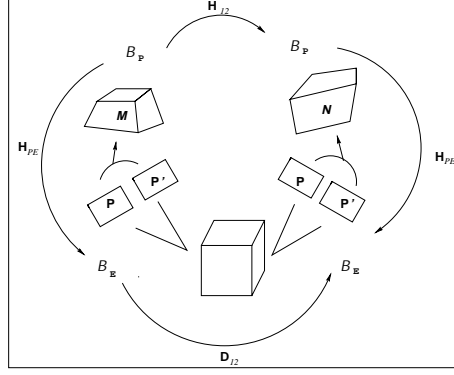
It is straightforward to show that, if  $\mathbf{H}_{PE}$  represents the collineation between the projective basis  $\mathcal{B}_P$  and the Euclidean frame  $\mathcal{B}_E$ , i.e.,  $\mathbf{M}_E \simeq \mathbf{H}_{PE}\mathbf{M}$ , we have the followings relations:

$$\mathbf{P} \simeq \mathbf{P}_E\mathbf{H}_{PE} \quad \text{and} \quad \mathbf{P}' \simeq \mathbf{P}'_E\mathbf{H}_{PE}$$

Indeed, from  $\mathbf{P}\mathbf{M} \simeq \mathbf{m} \simeq \mathbf{P}_E\mathbf{M}_E \simeq \mathbf{P}_E\mathbf{H}_{PE}\mathbf{M}$  it results that  $\mathbf{P} \simeq \mathbf{P}_E\mathbf{H}_{PE}$ .

The basic assumption throughout the paper is that the stereo rig performs a series of rigid motions and that during these motions  $\mathbf{K}, \mathbf{K}', \mathbf{R}$ , and  $\mathbf{t}$  remain constant, as shown in Figure 1. As the bases  $\mathcal{B}_P$  and  $\mathcal{B}_E$  are related to the

<sup>2</sup> We denote by  $\mathbf{I}_n$  the  $n \times n$  matrix of identity.



**Fig. 1.** Rigid motion of a stereo rig.

stereo rig we can again use them to compute  $\mathbf{N}$  and  $\mathbf{N}_E$ , the projective and Euclidean representations of the same physical point after the motion. Clearly, the relationship between the projective and Euclidean representations before the motion holds after the motion,  $\mathbf{N}_E \simeq \mathbf{H}_{PE}\mathbf{N}$ .

Let  $\mathbf{D}_{12}$  be the  $4 \times 4$  matrix describing the rigid motion performed by the stereo rig. We have  $\mathbf{N}_E = \mathbf{D}_{12}\mathbf{M}_E$  and by substituting  $\mathbf{M}_E$  and  $\mathbf{N}_E$  with  $\mathbf{H}_{PE}\mathbf{M}$  and  $\mathbf{H}_{PE}\mathbf{N}$  we obtain:

$$\mathbf{N} \simeq \mathbf{H}_{PE}^{-1}\mathbf{D}_{12}\mathbf{H}_{PE}\mathbf{M}$$

Consequently, the collineation between the two projective reconstructions (before and after the motion) ( $\mathbf{M}$  and  $\mathbf{N}$ ) is related to the rigid motion by the following formula:

$$\mathbf{H}_{12} \simeq \mathbf{H}_{PE}^{-1}\mathbf{D}_{12}\mathbf{H}_{PE} \quad (6)$$

In order to get rid of the scale factor one may normalize  $\mathbf{H}_{12}$  by dividing each term with  $\text{sign}(\text{trace}(\mathbf{H}_{12})) \sqrt[4]{\det(\mathbf{H}_{12})}$ . With this normalization “ $\simeq$ ” becomes “ $=$ ” and the eigenvalues of  $\mathbf{H}_{12}$  and of  $\mathbf{D}_{12}$  are the same. In what follows we assume that  $\mathbf{H}_{12}$  has been normalized.

The following proposition proved in [8] makes explicit the structure of  $\mathbf{H}_{PE}$  under the choice of the bases  $\mathcal{B}_P$  and  $\mathcal{B}_E$ :

**Proposition 1.** *The  $4 \times 4$  collineation  $\mathbf{H}_{PE}$  allowing the conversion of a projective reconstruction (basis  $\mathcal{B}_P$ ) obtained with a stereo rig into an Euclidean reconstruction (basis  $\mathcal{B}_E$ ) has the following structure:*

$$\mathbf{H}_{PE} = \begin{pmatrix} \mathbf{K}^{-1} & \mathbf{0}_3 \\ \mathbf{a}^\top & \lambda \end{pmatrix} \quad (7)$$

where  $\mathbf{K}$  is the matrix of intrinsic parameters of the left camera and  $(\mathbf{a}^\top \lambda)$  is the equation of the plane at infinity in the projective basis  $\mathcal{B}_P$ .

## 4 Algebraic preliminaries

In this section we make explicit some algebraic properties of the  $4 \times 4$  collineation  $\mathbf{H}_{12}$  which are direct consequences of equation (6).

The displacement matrix  $\mathbf{D}_{12}$  has the form:

$$\mathbf{D} = \begin{pmatrix} \mathbf{R}_{12} & \mathbf{t}_{12} \\ \mathbf{0}_3^\top & 1 \end{pmatrix} \quad (8)$$

where  $\mathbf{R}_{12}$  is a  $3 \times 3$  rotation matrix and  $\mathbf{t}_{12}$  represents the translation. Therefore the eigenvalues of  $\mathbf{D}_{12}$  are  $\{e^{i\theta}, e^{-i\theta}, 1, 1\}$ .

A key issue with our approach is the distinction between the *algebraic multiplicity* of an eigenvalue and its *geometric multiplicity*. We recall the following definitions (see [16,9]):

**Definition 2.** Let  $\lambda$  be an eigenvalue of a matrix. Its **algebraic multiplicity** is the number of times that it occurs as a root of the corresponding characteristic polynomial.

**Definition 3.** The **geometric multiplicity** of an eigenvalue  $\lambda$  is the dimension of the eigenspace associated with the eigenvalue  $\lambda$ .

In the case of a rigid displacement the algebraic multiplicity of the eigenvalue  $\lambda = 1$  is, in general, equal to 2 (it is equal to 4 if  $\theta = 0$ ). However its geometric multiplicity depends on whether the rigid motion is a screw or not and is equal to:

- 1 in the case of general displacement, i.e., there is a translation along the screw axis
- 2 in the case of planar motion (there is no translation along the axis or rotation).

Therefore there are distinct calibration solutions for these two types of motion.

Given matrix  $\mathbf{H}_{12}$  similar to  $\mathbf{D}_{12}$ , its eigenvalues are  $\{e^{i\theta}, e^{-i\theta}, 1, 1\}$  as well. From [9] we have the following proposition.

**Proposition 4.** Let  $\mathbf{\Omega}$  be a matrix with its eigenvalues equal to  $\{e^{i\theta}, e^{-i\theta}, 1, 1\}$ . Then there exists a non-singular matrix  $\mathbf{\Lambda}$  such that:

$$\mathbf{\Omega} = \mathbf{\Lambda} \begin{pmatrix} \cos(\theta) & -\sin(\theta) & 0 & 0 \\ \sin(\theta) & \cos(\theta) & 0 & 0 \\ 0 & 0 & 1 & \varepsilon \\ 0 & 0 & 0 & 1 \end{pmatrix} \mathbf{\Lambda}^{-1} = \mathbf{\Lambda} \mathbf{J}_\varepsilon \mathbf{\Lambda}^{-1} \quad (9)$$

where  $\varepsilon = 1$ , if the geometric multiplicity of the double eigenvalue  $\lambda = 1$  is 1,  $\mathbf{J}_\varepsilon = \mathbf{J}_1$ , and  $\varepsilon = 0$  if the geometric multiplicity of the double eigenvalue  $\lambda = 1$  is equal to 2,  $\mathbf{J}_\varepsilon = \mathbf{J}_0$ .

Furthermore, if  $\mathbf{\Omega} = \mathbf{D}$  is a displacement,  $\mathbf{\Lambda} = \mathbf{\Sigma}$  is of form:

$$\mathbf{\Sigma} = \begin{pmatrix} \mathbf{Q}_\Sigma & \mathbf{t}_\Sigma \\ \mathbf{0}_3^\top & 1 \end{pmatrix} \quad (10)$$

where  $\mathbf{Q}_\Sigma$  is an orthogonal matrix, i.e.  $\mathbf{\Sigma}$  is a rotation or a reflection followed by a translation.

Since  $\mathbf{J}_\varepsilon$  is the matrix associated with the *real Jordan canonical form*, the factorizations introduced in the proposition above are called *real Jordan factorizations*.

#### 4.1 Non uniqueness of the real Jordan factorization

The real Jordan factorization described above is not unique. This non unicity is of crucial importance for our calibration method and we are going to give some insights into this property.

**Proposition 5.** *The real Jordan factorization of a matrix  $\mathbf{\Omega}$  is not unique. Moreover,  $\mathbf{\Omega} = \mathbf{\Lambda}\mathbf{J}_\varepsilon\mathbf{\Lambda}^{-1}$  and  $\mathbf{\Omega} = \mathbf{\Lambda}'\mathbf{J}_\varepsilon\mathbf{\Lambda}'^{-1}$  are real Jordan factorizations if and only if there exists a matrix  $\mathbf{M}$  commuting with  $\mathbf{J}_\varepsilon$ , i.e.  $\mathbf{M}\mathbf{J}_\varepsilon = \mathbf{J}_\varepsilon\mathbf{M}$ , such that  $\mathbf{\Lambda}' = \mathbf{\Lambda}\mathbf{M}$ .*

**Proof:** For any  $\mathbf{M}$  commuting with  $\mathbf{J}_\varepsilon$ , we have  $\mathbf{\Omega} = \mathbf{\Lambda}\mathbf{J}_\varepsilon\mathbf{M}\mathbf{M}^{-1}\mathbf{\Lambda}^{-1} = \mathbf{\Lambda}\mathbf{M}\mathbf{J}_\varepsilon(\mathbf{\Lambda}\mathbf{M})^{-1}$ .

Conversely, if  $\mathbf{\Omega} = \mathbf{\Lambda}\mathbf{J}_\varepsilon\mathbf{\Lambda}^{-1} = \mathbf{\Lambda}'\mathbf{J}_\varepsilon\mathbf{\Lambda}'^{-1}$ , results that  $\mathbf{\Lambda}^{-1}\mathbf{\Lambda}'\mathbf{J}_\varepsilon = \mathbf{J}_\varepsilon(\mathbf{\Lambda}^{-1}\mathbf{\Lambda}')^{-1}$ . Hence  $\mathbf{\Lambda}^{-1}\mathbf{\Lambda}' = \mathbf{M}$  commutes with  $\mathbf{J}_\varepsilon$ . ■

By making explicit the matrix equality  $\mathbf{M}\mathbf{J}_\varepsilon = \mathbf{J}_\varepsilon\mathbf{M}$ , it is easy to derive the structure of  $\mathbf{M}$ :

- if  $\varepsilon = 1$  (general motion),  $\mathbf{M} = \mathbf{M}_g$  has 4 degrees of freedom and can be written as:

$$\mathbf{M}_g = \begin{pmatrix} \alpha - \beta & 0 & 0 \\ \beta & \alpha & 0 & 0 \\ 0 & 0 & \gamma & \omega \\ 0 & 0 & 0 & \gamma \end{pmatrix} \quad (11)$$

- if  $\varepsilon = 0$  (planar motion),  $\mathbf{M} = \mathbf{M}_p$  has 6 degrees of freedom and the form:

$$\mathbf{M}_p = \begin{pmatrix} \alpha - \beta & 0 & 0 \\ \beta & \alpha & 0 & 0 \\ 0 & 0 & \gamma & \omega \\ 0 & 0 & \delta & \eta \end{pmatrix} \quad (12)$$



– if  $\mathbf{\Omega}$  is a displacement, i.e.  $\mathbf{\Omega} = \mathbf{\Sigma}\mathbf{J}_\varepsilon\mathbf{\Sigma}^{-1}$ ,  $\mathbf{\Sigma}\mathbf{M}$ , must be of form (10). Using this constraint we obtain:

$$\mathbf{M}_d = \begin{pmatrix} \cos(\psi) & \sin(\psi) & 0 & 0 \\ -\sin(\psi) & \cos(\psi) & 0 & 0 \\ 0 & 0 & \pm 1 & \omega \\ 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{Q}_z & t_z \\ \mathbf{0}_3^\top & 1 \end{pmatrix} \quad (13)$$

## 4.2 The real Jordan factorization of a collineation

The real Jordan factorization will be used to decompose matrix  $\mathbf{H}_{12}$  which maps two projective reconstructions obtained with the stereo rig before and after the motion. It is therefore important to describe how to obtain **one** such a factorization.

**Proposition 6.** *Let  $\{e^{i\theta}, e^{-i\theta}, 1, 1\}$  be the eigenvalues of  $\mathbf{H}_{12}$  and let  $\mathbf{v}_1, \mathbf{v}_2$  be the eigenvectors associated with  $e^{i\theta}$  and  $e^{-i\theta}$ .*

(i) *If the eigenvalue  $\lambda = 1$  has geometric multiplicity equal to 1, let  $\mathbf{u}_3$  be the eigenvector associated with this eigenvalue and we obtain the following real Jordan factorization:*

$$\mathbf{H}_{12} = (\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3 \ \mathbf{w}) \mathbf{J}_1 (\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3 \ \mathbf{w})^{-1} = \mathbf{\Lambda}_g \mathbf{J}_1 \mathbf{\Lambda}_g^{-1} \quad (14)$$

where  $\mathbf{u}_1 = \mathbf{v}_1 + \mathbf{v}_2$ ,  $\mathbf{u}_2 = i(\mathbf{v}_1 - \mathbf{v}_2)$  are two real vectors and vector  $\mathbf{w}$  is defined by<sup>3</sup>

$$\mathbf{w} = \begin{pmatrix} (\overline{\mathbf{H}}_{12} - \mathbf{I}_3)^{-1} \overline{\mathbf{u}}_3 \\ w_4 \end{pmatrix} \quad (15)$$

with  $w_4$  chosen such that  $\det(\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{w}) \neq 0$ .

(ii) *If the eigenvalue  $\lambda = 1$  has geometric multiplicity equal to 2 let the vectors  $\{\mathbf{u}_3, \mathbf{u}_4\}$  be a basis of the associated eigenspace and in this case the following real Jordan factorization is obtained:*

$$\mathbf{H}_{12} = (\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3 \ \mathbf{u}_4) \mathbf{J}_0 (\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3 \ \mathbf{u}_4)^{-1} = \mathbf{\Lambda}_p \mathbf{J}_0 \mathbf{\Lambda}_p^{-1} \quad (16)$$

**Proof:** We show (ii) first. From  $\mathbf{H}_{12}\mathbf{v}_1 = (\cos(\theta) + i\sin(\theta))\mathbf{v}_1$  and  $\mathbf{H}_{12}\mathbf{v}_2 = (\cos(\theta) - i\sin(\theta))\mathbf{v}_2$ , by simple addition and subtraction results:

$$\begin{cases} \mathbf{H}_{12}\mathbf{u}_1 = \cos(\theta)\mathbf{u}_1 + \sin(\theta)\mathbf{u}_2 \\ \mathbf{H}_{12}\mathbf{u}_2 = -\sin(\theta)\mathbf{u}_1 + \cos(\theta)\mathbf{u}_2 \end{cases} \quad (17)$$

Furthermore  $\mathbf{H}_{12}\mathbf{u}_3 = \mathbf{u}_3$ ,  $\mathbf{H}_{12}\mathbf{u}_4 = \mathbf{u}_4$  and writing together with (17), using matrix notation, gives:

$$\mathbf{H}_{12} (\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3 \ \mathbf{u}_4) = (\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3 \ \mathbf{u}_4) \begin{pmatrix} \cos(\theta) - \sin(\theta) & 0 & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

<sup>3</sup> We denoted by  $\overline{\mathbf{A}}$  the  $3 \times 3$  upper left block of a  $4 \times 4$  matrix  $\mathbf{A}$  and by  $\overline{\mathbf{a}}$  the first 3 components of the 4 vector  $\mathbf{a}$ .

which is equivalent to (16).

Second we prove (i). By combining equation (17) with  $\mathbf{H}_{12}\mathbf{u}_3 = \mathbf{u}_3$  we obtain:

$$\mathbf{H}_{12}(\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3) = (\mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3) \begin{pmatrix} \cos(\theta) - \sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

By inspecting equation (14), we notice that  $\mathbf{w}$  must verify:

$$\mathbf{H}_{12}\mathbf{w} = \mathbf{u}_3 + \mathbf{w} \quad (18)$$

The rank of  $\mathbf{H}_{12} - \mathbf{I}_4$  is equal to 3. We can assume that  $\det(\overline{\mathbf{H}}_{12} - \mathbf{I}_3) \neq 0$  and construct the vector given by equation (15). If  $\det(\overline{\mathbf{H}}_{12} - \mathbf{I}_3) = 0$ , we can consider an other  $3 \times 3$  block of the matrix  $\mathbf{H}_{12}$  and apply the same method. ■

Consequently, one way to compute a real Jordan factorization of a  $4 \times 4$  collineation  $\mathbf{H}_{12}$  is to compute its eigenvalues and associated eigenvectors. In practice  $\mathbf{H}_{12}$  is estimated from image measurements and therefore it is corrupted by noise. This noise can considerably perturbate the eigenvalues and eigenvectors of  $\mathbf{H}_{12}$  [16].

Therefore we devised a simple method for computing the column vectors of matrix  $\mathbf{A}$  without computing explicitly the eigenvalues of  $\mathbf{H}_{12}$ . The latter is estimated up to a scale factor but after normalization one may notice that the trace has a simple form  $\text{trace}(\mathbf{H}_{12}) = 2 + 2\cos(\theta)$ . Therefore, we have  $\cos(\theta) = \frac{\text{trace}(\mathbf{H}) - 2}{2}$  and  $\sin(\theta) = \sqrt{1 - \cos(\theta)^2}$ . Finally, from equation (17) we obtain:

$$\begin{pmatrix} \mathbf{H}_{12} - \cos(\theta)\mathbf{I}_4 & -\sin(\theta)\mathbf{I}_4 \\ \sin(\theta)\mathbf{I}_4 & \mathbf{H}_{12} - \cos(\theta)\mathbf{I}_4 \end{pmatrix} \begin{pmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \end{pmatrix} = \mathbf{0}_8$$

which yields a solution for  $\mathbf{u}_1$  and  $\mathbf{u}_2$ .

The eigenspace corresponding to the eigenvalue 1 is given by  $(\mathbf{H}_{12} - \mathbf{I}_4)\mathbf{u} = \mathbf{0}_4$ . In the noise-less case the rank of  $\mathbf{H}_{12} - \mathbf{I}_4$  is equal to 3. However, when the data are corrupted by noise  $\det(\mathbf{H}_{12} - \mathbf{I}_4) \neq 0$  and an approximate solution must be found. In practice the singular value decomposition of  $\mathbf{H}_{12} - \mathbf{I}_4$  allows to compute the eigenspace associated with the unit eigenvalue. If there is one small singular value, the geometric multiplicity is one and  $\mathbf{u}_3$  is the vector corresponding to this singular value. If there are two small singular values, the geometric multiplicity is 2 and the two associated vectors are  $\mathbf{u}_3$  and  $\mathbf{u}_4$ .

## 5 Euclidean calibration

We come back now to the basic equation associated with the rigid motion of a stereo rig  $\mathbf{H}_{12} = \mathbf{H}_{PE}^{-1}\mathbf{D}_{12}\mathbf{H}_{PE}$ . Consider first a real Jordan factorization of

$\mathbf{D}_{12}$ , i.e.  $\mathbf{D}_{12} = \Sigma \mathbf{J} \Sigma^{-1}$ . We obtain **all** the factorization by multiplying  $\Sigma$  with  $\mathbf{M}_d$ , equation (13):

$$\mathbf{D}_{12} = \Sigma \mathbf{M}_d \mathbf{J}_\varepsilon (\Sigma \mathbf{M}_d)^{-1}$$

Consider now a real Jordan factorization of  $\mathbf{H}_{12}$  obtained as described in Section 4.2. Again we multiply by matrix  $\mathbf{M}$  to obtain all possible factorizations of  $\mathbf{H}_{12}$ , i.e.  $\mathbf{H}_{12} = \Lambda \mathbf{M} \mathbf{J}_\varepsilon (\Lambda \mathbf{M})^{-1}$ .

Replacing  $\mathbf{D}_{12}$  and  $\mathbf{H}_{12}$  in  $\mathbf{H}_{12} = \mathbf{H}_{PE}^{-1} \mathbf{D}_{12} \mathbf{H}_{PE}$ , results:

$$\Lambda \mathbf{M} \mathbf{J}_\varepsilon (\Lambda \mathbf{M})^{-1} = \mathbf{H}_{PE}^{-1} \Sigma \mathbf{M}_d \mathbf{J}_\varepsilon (\mathbf{H}_{PE}^{-1} \Sigma \mathbf{M}_d)^{-1}$$

We immediately obtain  $\mathbf{H}_{PE}^{-1} \Sigma \mathbf{M}_d = \Lambda \mathbf{M}$  and, from equations (7), (10) and (13), results:

$$\begin{pmatrix} \mathbf{K} & \mathbf{0}_3 \\ -\frac{1}{\lambda} \mathbf{a}^\top \mathbf{K} & \frac{1}{\lambda} \end{pmatrix} \begin{pmatrix} \mathbf{Q}_\Sigma t_\Sigma \\ \mathbf{0}_3^\top & 1 \end{pmatrix} \begin{pmatrix} \mathbf{Q}_z t_z \\ \mathbf{0}_3^\top & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{K} \mathbf{Q}_\Sigma \mathbf{Q}_z & \bullet \\ \bullet & \bullet \end{pmatrix}$$

Furthermore, by considering only the upper-left  $3 \times 3$  block matrices in this equation one obtains  $\mathbf{K} \mathbf{Q}_\Sigma \mathbf{Q}_z = \overline{(\Lambda \mathbf{M})}$  and finally the orthogonality of matrices  $\mathbf{Q}_\Sigma$  and  $\mathbf{Q}_z$  leads to the following relationship:

$$\mathbf{K} \mathbf{K}^\top = (\mathbf{K} \mathbf{Q}_\Sigma \mathbf{Q}_z) (\mathbf{K} \mathbf{Q}_\Sigma \mathbf{Q}_z)^\top = \overline{(\Lambda \mathbf{M})} \overline{(\Lambda \mathbf{M})}^\top \quad (19)$$

## 5.1 General motion

For a general motion the structure of matrix  $\mathbf{M}$  is given by equations (14) and (11) and we have:

$$\Lambda_g \mathbf{M}_g = (\alpha \mathbf{u}_1 + \beta \mathbf{u}_2 \ \alpha \mathbf{u}_2 - \beta \mathbf{u}_1 \ \gamma \mathbf{u}_3 \ \omega \mathbf{u}_3 + \gamma \mathbf{w})$$

The dual of the image of the absolute conic becomes in this case:

$$\mathbf{K} \mathbf{K}^\top = \underbrace{(\overline{\Lambda_g \mathbf{M}_g}) (\overline{\Lambda_g \mathbf{M}_g})^\top}_{\tau} = \underbrace{(\overline{\mathbf{u}_1} \overline{\mathbf{u}_1}^\top + \overline{\mathbf{u}_2} \overline{\mathbf{u}_2}^\top)}_{\sigma} + \underbrace{\gamma^2}_{\sigma} \overline{\mathbf{u}_3} \overline{\mathbf{u}_3}^\top \quad (20)$$

where  $\overline{\mathbf{u}_i} = (u_{i1}, u_{i2}, u_{i3})^\top$  for each  $\mathbf{u}_i = (u_{i1}, u_{i2}, u_{i3}, u_{i4})^\top$ ,  $i = 1, 2, 3$ .

Note that  $\mathbf{K} \mathbf{K}^\top$  depends merely on vectors  $\overline{\mathbf{u}_1}, \overline{\mathbf{u}_2}, \overline{\mathbf{u}_3}$  which have already been estimated and on two further positive parameters  $\tau \geq 0$  and  $\sigma \geq 0$ . Therefore, in order to estimate  $\mathbf{A} = \mathbf{K} \mathbf{K}^\top$  from a single movement two additional constraints are needed. As it was shown in Section 2, a four-parameter camera has exactly two constraints associated with the entries of  $\mathbf{A}$ , namely:

$$\begin{cases} A_{33} = 1 \\ A_{12} - A_{13} A_{23} = 0 \end{cases} \quad (21)$$

By combining (20) with (21) obtain the following solutions for  $\tau$  and  $\sigma$ :

$$\begin{aligned} \sigma &= -\frac{\prod_{j=1}^2 (\mathbf{u}_1^j \mathbf{u}_2^3 - \mathbf{u}_1^3 \mathbf{u}_2^j)}{\prod_{j=1}^2 (\mathbf{u}_3^3 (\mathbf{u}_1^j \mathbf{u}_1^3 + \mathbf{u}_2^j \mathbf{u}_2^3) - \mathbf{u}_3^j ((\mathbf{u}_1^3)^2 + (\mathbf{u}_2^3)^2))} \\ \tau &= \frac{1 - \sigma (\mathbf{u}_3^3)^2}{(\mathbf{u}_1^3)^2 + (\mathbf{u}_2^3)^2} \end{aligned} \quad (22)$$

where  $\mathbf{u}_i^j$  is the  $j^{\text{th}}$  component of  $\mathbf{u}_i$ . Notice that one must check the sign of  $\sigma$  and  $\tau$  since they must be, strictly positives by definition.

Once  $\sigma$  and  $\tau$  are computed, one may determine  $\mathbf{K}\mathbf{K}^\top$  and compute the intrinsic parameters either directly from equation (5) or by Cholesky factorization.

## 5.2 Planar motion

We consider now the case of a planar motion. In this case the structure of matrix  $\mathbf{M}$  is defined by equation (12) and the structure of matrix  $\mathbf{A}$  is given by equation (16), i.e., matrices  $\mathbf{M}_p$  and  $\mathbf{A}_p$ . Hence:

$$\mathbf{A}_p \mathbf{M}_p = (\alpha \mathbf{u}_1 + \beta \mathbf{u}_2 \quad \alpha \mathbf{u}_2 - \beta \mathbf{u}_1 \quad \gamma \mathbf{u}_3 + \delta \mathbf{u}_4 \quad \omega \mathbf{u}_3 + \eta \mathbf{u}_4)$$

and we obtain for the dual of the image of the absolute conic:

$$\begin{aligned} \mathbf{K}\mathbf{K}^\top &= \overline{(\mathbf{A}_p \mathbf{M}_p)} \overline{(\mathbf{A}_p \mathbf{M}_p)}^\top = \underbrace{(\alpha^2 + \beta^2)}_\tau (\bar{\mathbf{u}}_1 \bar{\mathbf{u}}_1^\top + \bar{\mathbf{u}}_2 \bar{\mathbf{u}}_2^\top) \\ &\quad + \underbrace{\gamma^2}_\sigma (\bar{\mathbf{u}}_3 \bar{\mathbf{u}}_3^\top + \underbrace{\left(\frac{\delta}{\gamma}\right)}_\mu (\bar{\mathbf{u}}_3 \bar{\mathbf{u}}_4^\top + \bar{\mathbf{u}}_4 \bar{\mathbf{u}}_3^\top) + \underbrace{\left(\frac{\delta^2}{\gamma^2}\right)}_{\mu^2} \bar{\mathbf{u}}_4 \bar{\mathbf{u}}_4^\top) \end{aligned} \quad (23)$$

In this case,  $\mathbf{A} = \mathbf{K}\mathbf{K}^\top$  is defined by vectors  $\bar{\mathbf{u}}_1, \bar{\mathbf{u}}_2, \bar{\mathbf{u}}_3, \bar{\mathbf{u}}_4$  and by three undetermined parameters  $\tau, \sigma$  and  $\mu$ . Since we have three unknown parameters, we need 3 constraints on  $\mathbf{A}$  in order to calibrate the camera with a single motion. A three-parameter camera has the following three constraints associated with it (see Section 2):

$$\begin{cases} A_{33} = 1 \\ A_{12} - A_{13}A_{23} = 0 \\ k^2(A_{11} - A_{12}^2) - (A_{23}^2 - A_{22}) = 0 \end{cases} \quad (24)$$

By combining (23) with (24) we obtain the following formulae for  $\tau > 0$ ,  $\sigma > 0$  and  $\mu \neq 0$ :

$$\begin{aligned} \mu &= -\frac{k^2 n^1 (\mathbf{u}_3^3 m^1 - \mathbf{u}_3^1 m^3) + n^2 (\mathbf{u}_3^3 m^2 - \mathbf{u}_3^2 m^3)}{k^2 n^1 (\mathbf{u}_4^3 m^1 - \mathbf{u}_4^1 m^3) + n^2 (\mathbf{u}_4^3 m^2 - \mathbf{u}_4^2 m^3)} \\ \sigma &= \left( \frac{k^2 m^3 n^1 \mathbf{u}_4^1 + m^3 n^2 \mathbf{u}_4^2 - \mathbf{u}_4^3 (m^1 n^1 k^2 + m^2 n^2)}{k m^3 (\mathbf{u}_3^3 (\mathbf{u}_4^3 m^1 - \mathbf{u}_4^1 m^3) - \mathbf{u}_3^2 (\mathbf{u}_4^3 m^2 - \mathbf{u}_4^2 m^3) + \mathbf{u}_3^3 (\mathbf{u}_4^2 m^1 - \mathbf{u}_4^1 m^2))} \right)^2 \\ \tau &= \frac{1 - \sigma (\mathbf{u}_3^3 + \mu \mathbf{u}_4^3)^2}{m^3} \end{aligned} \quad (25)$$

with  $j \in \{1, 2, 3\}$ ,  $m^j = \mathbf{u}_1^j \mathbf{u}_1^3 + \mathbf{u}_2^j \mathbf{u}_2^3$  and  $n^j = \mathbf{u}_1^j \mathbf{u}_2^3 - \mathbf{u}_1^3 \mathbf{u}_2^j$ .

Once  $\tau, \sigma$  and  $\mu$  are thus computed and if the constraint  $\tau > 0$  is verified, one can determine the three camera parameters either in closed-form or by Cholesky factorization.

## 6 Experimental results

In a first series of experiments, the above developed methods are evaluated on synthetic image data in order to quantitatively study the accuracy of calibration as a function of image noise, and the type of movements considered. A second experiment compares the results of self-calibration and standard off-line calibration on images of a particular calibration grid. The third experiment is to validate the use of our method for self-calibration of a stereo rig in an unknown real-world scene.

### 6.1 Synthetic data

Synthetic stereo images showing a scene of about 40 3-D points are generated for five different points of view. The internal parameters of the two cameras were kept fixed, such that projective reconstruction using [5] with the same projection matrices results in a representation of the scene in one and the same projective frame related to the stereo rig.

The different viewpoints are related by rigid motions of the rig and the conjugate collineations  $\mathbf{H}_{ii+1}$  from position  $i$  to  $i+1$  of projective space are estimated by the linear method presented in [8]. General screw motions and restricted planar motions are considered as camera motions, in order to comparatively study the performance of the respective methods of self-calibration.

The influence of image noise is evaluated by adding artificial Gaussian noise with standard deviation varying from 0.3 to 2 pixels. In order to obtain significant results, closed-form self-calibration is performed for each single movement in the sequence and the average value of each parameter over the sequence is considered. For each parameter we compute the relative error between the estimated values  $\alpha_u^*, k^*, u_0^*, v_0^*$  and the true values  $\alpha_u, k, u_0, v_0$ :

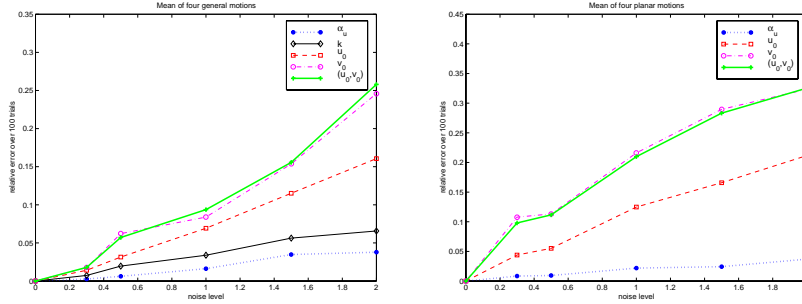
$$\epsilon_\alpha = \frac{|\alpha_u - \alpha_u^*|}{|\alpha_u|}, \quad \epsilon_k = \frac{|k - k^*|}{|k|}, \quad \epsilon_u = \frac{|u_0 - u_0^*|}{|u_0|} \quad \text{and} \quad \epsilon_v = \frac{|v_0 - v_0^*|}{|v_0|}$$

and for the principal point additionally the relative Euclidean distance between the points  $(u_0^*, v_0^*)$  and  $(u_0, v_0)$  is considered:

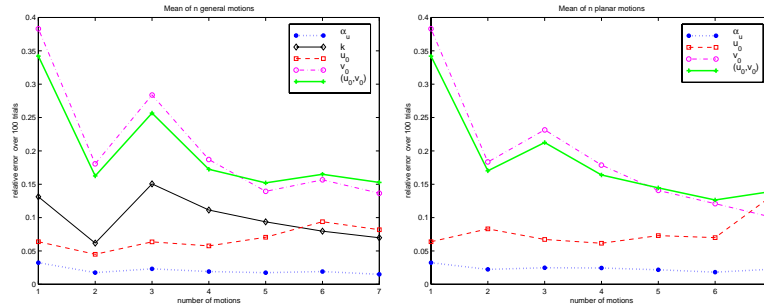
$$\epsilon_{u,v} = \sqrt{\frac{(u_0 - u_0^*)^2 + (v_0 - v_0^*)^2}{u_0^2 + v_0^2}}$$

The median of the relative errors over 100 trial runs is depicted in Figure 2 which demonstrates that accuracy of self-calibration degrades monotonically nicely with increasing measurement noise. Furthermore, calibration in the case of planar motion compares favorably with the case of general motion, given that a priori estimates of the fixed parameters (skew and aspect ratio) are in the vicinity of the true values.

Obviously, the estimates of the principal point  $(u_0, v_0)$  are less stable than those of the scale factors  $(\alpha_u, k)$ . On the one hand, the instability of the principal



**Fig. 2.** Relative error in estimates of the intrinsic parameter from four general displacements (left) and four planar motions (right) at different noise levels.



**Fig. 3.** The relative error in intrinsic parameters using an increasing number of motions.

point is an intrinsic problem of camera calibration from noisy image measurements and was observed with most of the existing algorithms. On the other hand, an inaccurate principal point barely affects Euclidean reconstruction, as outlined in [8].

To quantify the possible gain in accuracy from combining several motions, closed-form self-calibration is performed over one to seven movements out of trajectories consisting of either a general motion or a planar motion. From 100 trial at a noise level of 1 pixel, the first  $i$  movements of the trajectory were taken to calibrate and estimate the parameters by their average values. Figure 3 shows the evolution of the relative errors as a function of the number of motions which are considered.

Obviously, using several motions instead of a of a single motion does not improve the accuracy associated with the estimation of  $\alpha_u$  and barely improves the accuracy associated with  $u_0$ ,  $v_0$  or  $k$ .

Method	left camera				right camera			
	$\alpha$	$k$	$u_0$	$v_0$	$\alpha$	$k$	$u_0$	$v_0$
Off-line calibration	1534	.996	270	265	1520	.996	264	271
General motion	1550	.988	278	300	1533	.988	256	277
Planar motion	1570	$k^*$	261	291	1561	$k^*$	291	296

**Table 1.** Results for the left and right camera parameters using off-line calibration and self-calibration. The results shown are means of several motions. We used  $k^* = .996$  as known value for planar motions.

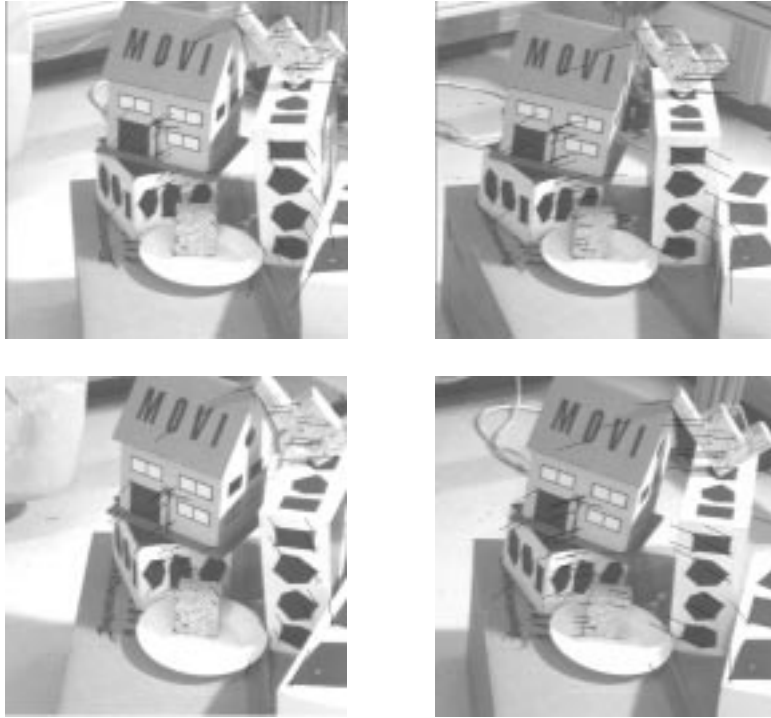
## 6.2 Self-calibration with image pairs of a calibration grid

To justify the applicability of our method for camera self-calibration and to compare its effectiveness with that of standard off-line methods, calibration is executed on images of a 3-D calibration grid. It consists of 100 circular target points, that are evenly distributed on three planes. Their 3-D positions are known with an accuracy of 0.02 mm, and their image projection are detected and localized at an accuracy of 0.05 pixel. The results of off-line calibration using [4] and the results of applying our self-calibration methods to eight stereo image pairs of the grid are compared in Table 1, for the left and right cameras, respectively. Even-though no knowledge about 3D scene structure is required, self-calibration performs, for both types of motions, as well as off-line calibration.

## 6.3 On-line self-calibration from image pairs of a 3-D scene

In order to validate the applicability of our method for camera self-calibration during runtime of a vision-system, i.e. for on-line calibration, we gathered 45 stereo images of a real-world scene from viewpoints which differ merely by small motions of the stereo pair. To obtain point correspondences we used the following stereo tracking algorithm. Interest points are extracted from the first pair and matched between the left and right images. Next, the points in the left image are tracked, as the stereo rig moves, over the sequence of images associated with the left camera motion. To obtain point matches associated with the image pair after the motion, the tracking is guided by the epipolar geometry. This algorithm makes use of the fact that the epipolar geometry remains unchanged during the motion of the stereo rig. Figure 4 shows the matches obtained with two image pairs.

In contrast to the previous experiments, the matched points are no longer evenly distributed, neither in the image, nor in space. Even worse, mismatches may be present in the data. Additionally, interest points are no longer extracted with 0.05 pixel accuracy. Nevertheless, the camera parameters resulting from self-calibration are within the expected range and the behaviour of the method is consistent with the results obtained for synthetic data. Figure 5 depicts the distribution of parameter estimates obtained by closed-form calibration for about 1000 planar or general motions. Table 2 shows the average values of each parameter set.



**Fig. 4.** Two image pairs and their matched points.

## 7 Conclusion

We proposed a method for self-calibration of a stereo rig from a single general or ground-plane motion. The basic assumption is that the stereo rig has the same internal and external parameters before and after the motion. In this case the collineation relating the two projective reconstructions performed at each position is conjugated to a displacement. By making explicit the algebraic properties of such a collineation, we derived a closed-form solution to recover four camera parameters from a single general motion or three camera parameters from a single ground-plane motion.

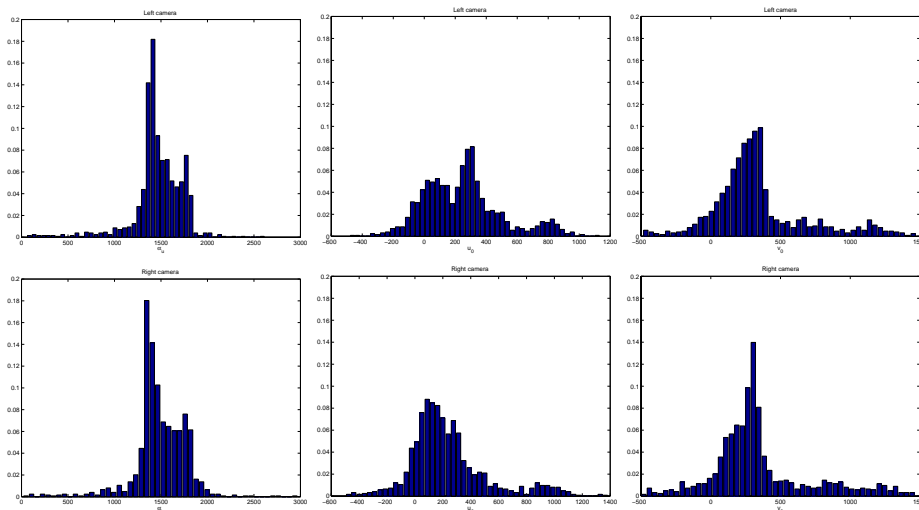
One of the advantages of this approach is that calibration was done without explicitly computing affine structure. The second advantage is that all the computations are based on linear algebra techniques such as singular value decomposition, and hence it is easy to implement and has low computation cost. Therefore, it can be included in on-line perception-action cycle, such as visual servoing.

One remarkable feature of our method is that it performs as well with small motions as with large ones. This has a crucial practical importance because it is much easier to find matches between images that differ by a small motion



Method	left camera				right camera			
	$\alpha$	$k$	$u_0$	$v_0$	$\alpha$	$k$	$u_0$	$v_0$
Off-line calibration	1534	.996	270	265	1520	.996	264	271
General motion	1531	1.01	255	323	1508	1.05	211	334
Planar motion	1462	$k^*$	154	246	1464	$k^*$	143	259

**Table 2.** Results for the left and right camera parameters using off-line calibration and self-calibration. The results shown are means of several motions. We used  $k^* = .996$  as known value for planar motions.



**Fig. 5.** The distributions of the estimated intrinsic parameters.

than for images that are far apart. One inconvenience is, however, that small motions may sometimes give rise to bad results simply because a rotation matrix which is closed to the identity matrix does not have the algebraic properties that are expected by the method. Nevertheless, these “bad” motions can be easily eliminated by simply checking the conditioning of the collineation  $\mathbf{H}_{12}$ .

Another source of errors is the sensitivity of the method to the accuracy with which the interest points are located and matched as well as the 3-D distribution of these points. This problem has been observed with both synthetic and real data. Interesting enough, the matrix conditioning analysis outlined above works well to eliminate such badly distributed data.

In the future we plan to investigate more thoroughly the relationship between bad calibration and the numerical conditioning of the problem and to combine closed-form calibration methods with statistically robust methods.

## References

1. P. A. Beardsley and A. Zisserman. Affine calibration of mobile vehicles. In *Proceedings of Europe-China Workshop on Geometrical Modelling and Invariants for Computer Vision*, pages 214–221, Xi'an, China, April 1995. Xidan University Press.
2. P.A. Beardsley, I.D. Reid, A. Zisserman, and D.W. Murray. Active visual navigation using non-metric structure. In E. Grimson, editor, *Proceedings of the 5th International Conference on Computer Vision, Cambridge, Massachusetts, USA*, pages 58–64. IEEE Computer Society Press, June 1995.
3. F. Devernay and O. Faugeras. From projective to Euclidean reconstruction. In *Proceedings Computer Vision and Pattern Recognition Conference*, pages 264–269, San Francisco, CA., June 1996.
4. O. D. Faugeras. *Three Dimensional Computer Vision: A Geometric Viewpoint*. MIT Press, Boston, 1993.
5. R. Hartley and P. Sturm. Triangulation. *Computer Vision and Image Understanding*, 68(2):146–157, 1997.
6. R. I. Hartley. Euclidean reconstruction from uncalibrated views. In Mundy Zisserman Forsyth, editor, *Applications of Invariance in Computer Vision*, pages 237–256. Springer Verlag, Berlin Heidelberg, 1994.
7. A. Heyden and K. Åström. Euclidean reconstruction from constant intrinsic parameters. In *Proceedings of the 13th International Conference on Pattern Recognition, Vienna, Austria*, volume I, pages 339–343, August 1996.
8. R. Horaud and G. Csurka. Self-calibration and euclidean reconstruction using motions of a stereo rig. In *Proceedings of the 6th International Conference on Computer Vision, Bombay, India*, pages 96–103, January 1998.
9. R.A. Horn and C.R. Johnson. *Matrix analysis*. Cambridge University Press, 1985.
10. Q-T. Luong and O. D. Faugeras. Self-calibration of a moving camera from point correspondences and fundamental matrices. *International Journal of Computer Vision*, 22(3):261–289, 1997.
11. Q.T. Luong and T. Vieville. Canonic representations for the geometries of multiple projective views. In *Proceedings of the 3rd European Conference on Computer Vision, Stockholm, Sweden*, pages 589–599, May 1994.
12. T. Moons, L. Van Gool, M. Proesmans, and E Pauwels. Affine reconstruction from perspective image pairs with a relative object-camera translation in between. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(1):77–83, January 1996.
13. M. Pollefeys and L. Van Gool. A stratified approach to metric self-calibration. In *Proceedings of the Conference on Computer Vision and Pattern Recognition, Puerto Rico, USA*, pages 407–412. IEEE Computer Society Press, June 1997.
14. P. Sturm. Critical motion sequences for monocular self-calibration and uncalibrated euclidean reconstruction. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1100–1105, Puerto-Rico, June 1997.
15. B. Triggs. Autocalibration and the absolute quadric. In *Proceedings of the Conference on Computer Vision and Pattern Recognition, Puerto Rico, USA*, pages 609–614. IEEE Computer Society Press, June 1997.
16. J.H. Wilkinson. *The algebraic eigenvalue problem*. Clarendon Press - Oxford, 1965.
17. A. Zisserman, P. A. Beardsley, and I. D. Reid. Metric calibration of a stereo rig. In *Proc. IEEE Workshop on Representation of Visual Scenes*, pages 93–100, Cambridge, Mass., June 1995.