



HAL
open science

Self-Calibration and Euclidean Reconstruction Using Motions of A Stereo Rig

Radu Horaud, Gabriela Csurka

► **To cite this version:**

Radu Horaud, Gabriela Csurka. Self-Calibration and Euclidean Reconstruction Using Motions of A Stereo Rig. 6th International Conference on Computer Vision (ICCV '98), Jan 1998, Bombay, India. pp.96–103, 10.1109/ICCV.1998.710706 . inria-00590095

HAL Id: inria-00590095

<https://inria.hal.science/inria-00590095>

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.

Self-calibration and Euclidean Reconstruction Using Motions of a Stereo Rig

Radu Horaud and Gabriella Csurka
GRAVIR-IMAG & INRIA Rhône-Alpes
655, avenue de l'Europe
38330 Montbonnot Saint Martin, FRANCE
E-mail: Radu.Horaud@inrialpes.fr

Abstract

This paper describes a method to upgrade projective reconstruction to affine and to metric reconstructions using rigid general motions of a stereo rig. We make clear the algebraic relationships between projective reconstruction, the plane at infinity (affine reconstruction), camera calibration, and metric reconstruction. We show that all the computations can be carried out using standard linear resolution methods and that these methods compare favorably with non-linear optimization methods in the presence of Gaussian noise. We carry out a theoretical error analysis which quantify the relative importance of the accuracies of projective-to-affine conversion and affine-to-Euclidean conversion. Experiments with real data are consistent with the theoretical error analysis and with a sensitivity analysis performed with simulated data.

What is the original contribution of this work? This paper describes a method for converting projective structure into affine and metric structure using *exclusively* linear resolution methods. We describe a novel method for estimating collineations between projective reconstructions. We show, based on algebraic properties, that the plane of infinity is the common eigenvector of several collineations associated with general rigid motion. We describe a theoretical error analysis in order to stratify the relative importance of affine and metric calibrations (or, equivalently, reconstructions).

Why should this contribution be considered important? The recovery of the Euclidean structure of sets of 3-D points without any sensor calibration is an important topic in itself and has a large number of applications.

What is the most closely related work by others and how does this work differ?

The most closely related works are by Zisserman et. al [16], Devernay and Faugeras [3], and Beardsley and Zisserman [2]. In [16] and [2] it is shown how to recover affine structure from a single general motion or from several ground-plane motions. We show, based on algebraic facts, how to combine several general motions in order to accurately compute the plane of infinity. In [16] metric structure is obtained from three fixed entities which appear as three virtual image points. We take a different approach and derive a more direct method for computing the infinite homography associated with motions of the left camera. In [3] a lower triangular collineation is introduced for upgrading projective to metric structure. We show that the algebraic structure of this conversion can be modified such that camera calibration appears explicitly. The advantage of this new formulation is that one can experiment with cameras with a variable number of parameters by fixing in advance some of the intrinsic parameters. None of the cited papers have done an extensive error and noise sensitivity analysis.

How can other researchers make use of the results of this work? The method's implementation is straightforward and is mainly based on such a standard technique as singular value decomposition of matrices. The error and sensitivity analysis allows one to predict the correctness of the result.

1 Introduction, background, and approach

In this paper we address the following problem: An uncalibrated stereo rig observes an unknown 3-D scene while it performs a set of rigid motions. A 3-D Euclidean reconstruction of the scene is desired. In the general case, 3-D structure can be recovered only up to a 3-D projective transformation. However, if the stereo rig undergoes a general motion and for unchanging intrinsic camera parameters, the projective ambiguity can be reduced to affine or to Euclidean. It is well known that the process of converting a projective reconstruction into an Euclidean one is equivalent to camera or stereo calibration.

The relationship between projective space, affine space, metric space and camera calibration has been thoroughly investigated both in the case of a moving unique camera and of a moving stereo rig. The Kruppa equations [13], [4], [10], [9] consists of a system of polynomial equations relating the intrinsic camera parameters to the epipolar geometry between two views taken with the camera. However, solving the Kruppa equations requires non-linear resolution methods. An alternative is to first recover affine structure and second solve for camera calibration using the affine calibration. This stratified approach [5] can be applied to a single camera in motion [6], [12] or to a stereo rig in motion [16], [3].

Affine calibration amounts to recover the position of the plane at infinity or, equivalently, the infinite homography between two views. In practice this is done using (i) *special camera motions* such as pure translations of a stereo rig [14], rotations around the camera's center of projection [7], or planar motions [2], [1], (ii) *exploiting special scene structure* such as parallel lines, or (iii) *using fixed entities* under rigid motion [16].

In this paper we investigate linear algebraic methods for recovering metric structure, affine calibration, and intrinsic camera parameters with an uncalibrated stereo rig, by performing a set of general rigid motions. 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 4×4 collineation \mathbf{H}_{12} which is related to the rigid motion \mathbf{D}_{12} by ([16], [3]):

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

where \mathbf{H}_{PE} is a 4×4 collineation allowing the projective reconstruction to be upgraded to Euclidean. It will be shown that this collineation encapsulates affine calibration of the

stereo rig and the intrinsic parameters of the left camera. If a 3-D point has projective coordinates $\mathbf{M}_1 \in \mathcal{P}_1$ and $\mathbf{M}_2 \in \mathcal{P}_2$ then $\mathbf{M}_2 = \mathbf{H}_{12}\mathbf{M}_1$. The Euclidean coordinates of the same point are $\mathbf{N}_1 = \mathbf{H}_{PE}\mathbf{M}_1$ and $\mathbf{N}_2 = \mathbf{H}_{PE}\mathbf{M}_2$ with $\mathbf{N}_2 = \mathbf{D}_{12}\mathbf{N}_1$.

Zisserman et al. [16] showed that the plane at infinity can be recovered from one eigenvector of matrix \mathbf{H}_{12}^{-T} and that intrinsic parameters of the left or right cameras can be recovered from three *virtual image points* that are fixed under Euclidean motion. Indeed the infinite homography between the images of the left camera (before and after the motion) can be determined from three image point correspondences if the two epipoles associated with this motion are known. The intrinsic parameters of the left camera (denoted by a 3×3 upper triangular matrix \mathbf{K}) are related to the infinite homography associated with the left camera motion, \mathbf{G}_{12} , and to the left camera rotation matrix \mathbf{R}_{12} by the following equation [12]:

$$\mathbf{G}_{12} = \mathbf{K}\mathbf{R}_{12}\mathbf{K}^{-1} \quad (2)$$

Using the orthogonality of rotation matrices one can easily obtain the following relationship:

$$\mathbf{G}_{12}^T\mathbf{K}^{-T}\mathbf{K}^{-1}\mathbf{G}_{12} = \mathbf{K}^{-T}\mathbf{K}^{-1} \quad (3)$$

Matrix $\mathbf{K}^{-T}\mathbf{K}^{-1}$ is known as the image of the absolute conic. Therefore, one can compute the image of the absolute conic, and hence the camera intrinsic parameters, if matrix \mathbf{G}_{12} is known.

We implemented the method described by Zisserman et al. and found that both the affine calibration and metric calibration methods are quite sensitive to noise. The reason for which affine calibration has higher noise sensitivity than expected is because the noise affects the algebraic properties of the collineation \mathbf{H}_{12} which are explicitly used for estimating the plane at infinity. Metric calibration relies on four image point correspondences – the epipole and three virtual points – the smallest number of matches needed to compute an image-to-image collineation (\mathbf{G}_{12}). A larger number of point matches is usually necessary to properly estimate the infinite homography. Moreover, the estimation of the two epipoles is very noise sensitive [11].

Devernay and Faugeras [3] showed that one possible factorization of \mathbf{H}_{12} in eq. (1) is such that \mathbf{H}_{PE} becomes a lower triangular matrix and the fourth row vector of this matrix is the plane at infinity. The authors propose a non-linear minimization method to

directly estimate Euclidean upgrading, i.e., the entries of \mathbf{H}_{PE} , from point correspondences between two stereo image pairs (before and after the motion). The method of Devernay and Faugeras gives interesting algebraic insights, although the algebraic properties associated with \mathbf{H}_{PE} are not used in practice. Moreover, the intrinsic camera parameters do not appear explicitly. In practice it is sometimes useful to assume that some of the intrinsic camera parameters are known (such as the image skew which is generally negligible) but this constraint cannot be used with this approach.

The method described in this article has the following contributions. First we show that with an appropriate choice for the Cartesian reference frame undergoing the rigid motion, the matrix \mathbf{H}_{PE} is parameterized by the plane at infinity and by the intrinsic parameters of the left camera. So, the homography \mathbf{H}_{PE} in eq. (1) directly encapsulates projective to Euclidean upgrading, affine calibration and left-camera calibration. This particular parameterization of \mathbf{H}_{PE} allows for an error analysis which determines the relative importance of affine calibration and metric calibration as well as the relative importance of the various intrinsic camera parameters.

The plane at infinity is an eigenvector of \mathbf{H}_{12}^{-T} or of \mathbf{H}_{12}^T . Second we show that for all rigid motions the corresponding collineations *have the same eigenvector*, otherwise this eigenvector is an intrinsic property of the stereo rig. This property allows us to estimate this eigenvector from any number of motions, the eigenvector being the common root of a set of linear equations. Once this eigenvector (the plane at infinity) has been recovered, the parameterization of \mathbf{H}_{12} in terms of \mathbf{H}_{PE} and \mathbf{D}_{12} provides a simple algebraic expression for \mathbf{G}_{12} , the infinite homography between the images associated with the left camera before and after a motion. This means that, unlike the Kruppa equations and unlike the method described in [16] it is not necessary to determine the epipolar geometry associated with the left (or right) camera motion.

The whole method described in this article heavily relies on the computation of collineations between two projective reconstructions. Third, we describe a novel linear method for estimating this collineation and we compare it experimentally with a non-linear least-square minimization method. We show (experimentally) that in the presence of Gaussian noise the linear method behaves as well as the non-linear one.

Finally, we describe experiments with both simulated and real data. The noise sensitiv-

ity analysis performed with simulated data allows to determine the optimal experimental conditions under which the method is expected to yield reliable camera calibration and metric reconstruction. The experiments performed with real data are consistent with this noise sensitivity analysis.

1.1 Paper organization

The remainder of the paper is organized as follows. Section 2 briefly recalls the classical geometric model associated with a pinhole camera, the relationship between camera calibration and the image of the absolute conic, and the geometry of a stereo rig. Section 3 describes the algebraic structure of the upgrading of projective structure to metric structure for the special case of a stereo rig undergoing rigid general motions. Section 4 describes and evaluates a linear method for estimating the 4×4 transformation between two projective reconstructions. Section 5 describes the method implementation and evaluates it with both simulated and real data. Finally, section 6 discusses the method in the light of the experimental results.

2 Preliminaries

2.1 Camera model and the image of 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 a 3×4 homogeneous matrix \mathbf{P} of rank equal to 3:

$$\mathbf{m} = \mathbf{P}\mathbf{M} \quad (4)$$

The equal sign 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 (the origin and orientation of the Euclidean frame is arbitrarily chosen):

$$\mathbf{P}_E = \mathbf{K} \begin{pmatrix} \mathbf{R} & \mathbf{t} \end{pmatrix} = \begin{pmatrix} \mathbf{KR} & \mathbf{Kt} \end{pmatrix} \quad (5)$$

If we choose 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 rotation matrix \mathbf{R} is equal to the identity matrix and the translation vector \mathbf{t} is the null vector. The projection matrix becomes:

$$\mathbf{P}_E = \begin{pmatrix} \mathbf{K} & \mathbf{0} \end{pmatrix} \quad (6)$$

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

$$\mathbf{K} = \begin{pmatrix} \alpha & r & u_0 \\ 0 & k\alpha & v_0 \\ 0 & 0 & 1 \end{pmatrix} \quad (7)$$

where α 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.

Eq. (7) 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 $r = 0$ which means that the image is a rectangle – a sensible assumption, or
- *three-parameter camera* with $r = 0$ and k having a known value; for instance the value of k can be obtained from the physical size of a pixel.

Let us make explicit the image of the absolute conic, i.e., the matrix $\mathbf{A} = \mathbf{K}^{-T}\mathbf{K}^{-1}$, since this matrix can be linearly estimated from eq. (3). In the general case (five-parameter camera) \mathbf{A} is a homogeneous, symmetric, semi-definite positive matrix.

With the constraint $r = 0$ we obtain:

$$\mathbf{A} = \mathbf{K}^{-T}\mathbf{K}^{-1} = \begin{pmatrix} k^2 & 0 & -k^2u_0 \\ 0 & 1 & -v_0 \\ -k^2u_0 & -v_0 & u_0^2k^2 + v_0^2 + \alpha^2k^2 \end{pmatrix} \quad (8)$$

With $r = 0$ and if k is known we can eliminate k and with the substitution $v'_0 = v_0/k$ we have:

$$\mathbf{A} = \mathbf{K}^{-T}\mathbf{K}^{-1} = \begin{pmatrix} 1 & 0 & -u_0 \\ 0 & 1 & -v'_0 \\ -u_0 & -v'_0 & u_0^2 + v'^2_0 + \alpha^2 \end{pmatrix} \quad (9)$$

2.2 The geometry 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 write these 3×4 matrices as:

$$\mathbf{P} = \begin{pmatrix} \bar{\mathbf{P}} & \mathbf{p} \end{pmatrix}$$

$$\mathbf{P}' = \begin{pmatrix} \bar{\mathbf{P}}' & \mathbf{p}' \end{pmatrix}$$

It is useful to recall the expressions of the infinite homography between the left and right images as well as the left and right epipoles:

$$\mathbf{H}_\infty = \bar{\mathbf{P}}' \bar{\mathbf{P}}^{-1} \quad (10)$$

and

$$\mathbf{e} = -\mathbf{H}_\infty^{-1} \mathbf{p}' + \mathbf{p} \quad (11)$$

$$\mathbf{e}' = -\mathbf{H}_\infty \mathbf{p} + \mathbf{p}' \quad (12)$$

In the uncalibrated case and without loss of generality the two projection matrices can be written as:

$$\mathbf{P} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \end{pmatrix} \quad (13)$$

$$\mathbf{P}' = \begin{pmatrix} \bar{\mathbf{P}}' & \mathbf{p}' \end{pmatrix} \quad (14)$$

In the calibrated (Euclidean) case one can use the following projection matrices (\mathbf{K}' is the matrix of right camera intrinsic parameters and \mathbf{R} and \mathbf{t} describe the orientation and position of the right camera frame with respect to the left camera frame):

$$\mathbf{P}_E = \begin{pmatrix} \mathbf{K} & \mathbf{0} \end{pmatrix}$$

$$\mathbf{P}'_E = \begin{pmatrix} \mathbf{K}'\mathbf{R} & \mathbf{K}'\mathbf{t} \end{pmatrix} \quad (15)$$

With these expressions for \mathbf{P} and \mathbf{P}' we obtain:

$$\mathbf{H}_\infty = \mathbf{K}'\mathbf{R}\mathbf{K}^{-1} \quad (16)$$

$$\mathbf{e} = -\mathbf{K}\mathbf{R}^T\mathbf{t} \quad (17)$$

$$\mathbf{e}' = \mathbf{K}'\mathbf{t} \quad (18)$$

2.3 Projective reconstruction with a stereo rig

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} from the equations $\mathbf{m} = \mu\mathbf{P}\mathbf{M}$ and $\mathbf{m}' = \mu'\mathbf{P}'\mathbf{M}$, where \mathbf{m} and \mathbf{m}' are the projections of \mathbf{M} onto the left and right images and μ and μ' are two unknown scale factors.

Matrices \mathbf{P} and \mathbf{P}' can be estimated from point matches without any camera calibration: Indeed, given at least 8 left-right image point correspondences, one can estimate the fundamental matrix which encapsulates the epipolar geometry for a pair of uncalibrated views [15], [8]. Several authors proved that the two projection matrices can be obtained from the epipolar geometry up to a 4-parameter projective mapping [12]:

$$\begin{aligned}\mathbf{P} &= \begin{pmatrix} \mathbf{I} & \mathbf{0} \end{pmatrix} \\ \mathbf{P}' &= \begin{pmatrix} \mathbf{H}_\infty + \mathbf{e}'\mathbf{a}^T & a\mathbf{e}' \end{pmatrix}\end{aligned}\tag{19}$$

where \mathbf{H}_∞ and \mathbf{e}' were defined above, \mathbf{a} is an arbitrary 3-vector and a is an arbitrary scale factor. It will be shown below that the 4-vector $(\mathbf{a}^T a)$ has a simple but important geometric interpretation.

3 From projective to metric reconstruction

We are interested into the problem of converting the 3-D projective reconstruction outlined above into a metric reconstruction. This conversion is a projective mapping from the projective space onto its Euclidean sub-space and this mapping is the 4×4 collineation \mathbf{H}_{PE} which appears in eq. (1). The left and right camera projections equations can be written as:

$$\mathbf{m} = \mathbf{P}\mathbf{H}_{PE}^{-1}\mathbf{H}_{PE}\mathbf{M}\tag{20}$$

$$\mathbf{m}' = \mathbf{P}'\mathbf{H}_{PE}^{-1}\mathbf{H}_{PE}\mathbf{M}\tag{21}$$

Since $\mathbf{N} = \mathbf{H}_{PE}\mathbf{M}$ is an Euclidean representation of \mathbf{M} , the projection matrices $\mathbf{P}\mathbf{H}_{PE}$ and $\mathbf{P}'\mathbf{H}_{PE}$ must have the structure given by eqs. (6) and (15). We can now state the following proposition:

Proposition 1 *The 4×4 collineation allowing the conversion of a projective reconstruction obtained with a stereo rig into an Euclidean reconstruction has the following structure:*

$$\mathbf{H}_{PE} = \begin{pmatrix} \mathbf{K}^{-1} & \mathbf{0} \\ \mathbf{a}^T & a \end{pmatrix} \quad (22)$$

where \mathbf{K} is the matrix of intrinsic parameters of the left camera and $(\mathbf{a}^T \ a)$ is the equation of the plane at infinity in the projective basis chosen to represent the projective reconstruction.

Indeed, the projection matrix of the left camera can be written as the following product:

$$\mathbf{P} = (\mathbf{I} \ \mathbf{0}) = (\mathbf{K} \ \mathbf{0}) \begin{pmatrix} \mathbf{K}^{-1} & \mathbf{0} \\ \mathbf{a}^T & a \end{pmatrix} = \mathbf{P}_E \mathbf{H}_e$$

By substituting eq. (16) and eq. (18) into eq. (19) we obtain:

$$\begin{aligned} \mathbf{P}' &= (\mathbf{H}_\infty + \mathbf{e}' \mathbf{a}^T \ a \mathbf{e}') \\ &= (\mathbf{K}' \mathbf{R} \mathbf{K}^{-1} + \mathbf{K}' \mathbf{t} \mathbf{a}^T \ a \mathbf{K}' \mathbf{t}) \\ &= (\mathbf{K}' \mathbf{R} \ \mathbf{K}' \mathbf{t}) \begin{pmatrix} \mathbf{K}^{-1} & \mathbf{0} \\ \mathbf{a}^T & a \end{pmatrix} \\ &= \mathbf{P}'_E \mathbf{H}_e \end{aligned}$$

Eqs. (20) and (21) become:

$$\mathbf{m} = \underbrace{\mathbf{P}_E \mathbf{H}_e}_{\mathbf{P}} \mathbf{H}_{PE}^{-1} \mathbf{H}_{PE} \mathbf{M}$$

and

$$\mathbf{m}' = \underbrace{\mathbf{P}'_E \mathbf{H}_e}_{\mathbf{P}'} \mathbf{H}_{PE}^{-1} \mathbf{H}_{PE} \mathbf{M}$$

By simply taking $\mathbf{H}_{PE} = \mathbf{H}_e$ proves the first part of the proposition.

In order to prove the second part of proposition 1 let us consider again the conjugate relationship of eq. (1). We immediately obtain:

$$\mathbf{H}_{12}^{-T} = \mathbf{H}_{PE}^T \mathbf{D}_{12}^{-T} \mathbf{H}_{PE}^{-T} \quad (23)$$

Since \mathbf{H}_{12} and \mathbf{D}_{12} are point transformation matrices, \mathbf{H}_{12}^{-T} and \mathbf{D}_{12}^{-T} are plane transformation matrices. Indeed, let $\mathbf{q}^T \mathbf{M} = 0$ be a plane of the 3-D projective space. By change of projective basis \mathbf{H} , point \mathbf{M} is mapped onto $\mathbf{M}' = \mathbf{H} \mathbf{M}$ and the plane equation in the new projective basis is $\mathbf{q}'^T \mathbf{M}' = 0$. By substitution we have $\mathbf{q}'^T \mathbf{H} \mathbf{M} = 0$ and by identification we obtain that $\mathbf{q}' = \mathbf{H}^{-T} \mathbf{q}$.

Matrix \mathbf{D}_{12} represents a rigid motion and hence its eigenvalues are $\lambda \in \{e^{i\theta}, e^{-i\theta}, 1, 1\}$. Therefore the eigenvalues of \mathbf{D}_{12}^{-T} are $1/\lambda \in \{e^{-i\theta}, e^{i\theta}, 1, 1\}$. The eigenvector associated with the double eigenvalue 1, $\mathbf{D}_{12}^{-T}\mathbf{y} = \mathbf{y}$, is obviously $\mathbf{y} = (0 \ 0 \ 0 \ 1)^T$ which is the plane of infinity in metric space. From eq. (23) we have that the eigenvector associated with the double eigenvalue 1 of \mathbf{H}_{12}^{-T} is the vector $\mathbf{H}_{PE}^T\mathbf{y}$ and we obtain:

$$\mathbf{H}_{12}^{-T} \mathbf{H}_{PE}^T \mathbf{y} = \mathbf{H}_{PE}^T \mathbf{y} = \begin{pmatrix} \mathbf{K}^{-T} & \mathbf{a} \\ \mathbf{0}^T & a \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \mathbf{a} \\ a \end{pmatrix}$$

This proves the second part of proposition 1.

3.1 Affine calibration

We can now derive another result which allows for affine calibration of the stereo rig:

Corollary 1.1 *Let us consider several **general** motions of the stereo rig, $\mathbf{D}_{12}, \mathbf{D}_{23}, \dots, \mathbf{D}_{n \ n+1}$. The matrices $\mathbf{H}_{12}^{-T}, \mathbf{H}_{23}^{-T}, \mathbf{H}_{n \ n+1}^{-T}$ (or, equivalently the matrices \mathbf{H}_{12}^T , etc.) with $\det \mathbf{H}_{ij} = 1$ have the same eigenvector associated with the double eigenvalue 1. This eigenvector is solution of the following set of linear homogeneous equations:*

$$\begin{pmatrix} \mathbf{H}_{12}^T - \mathbf{I} \\ \vdots \\ \mathbf{H}_{n \ n+1}^T - \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{a} \\ a \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \quad (24)$$

Indeed, from the above derivation it is straightforward to notice that the eigenvector of \mathbf{H}_{12}^T associated with the unit eigenvalue is not a function of the rigid motion \mathbf{D}_{12} of the stereo rig. Hence this eigenvector can be estimated as the common root of equations $(\mathbf{H}_{ij}^T - \mathbf{I})\mathbf{x} = 0$ for all motions from position i to position j .

We denote by \mathbf{B} the $4n \times 4$ (n being the number of motions) matrix present in eq. (24). In the noise-less case, the rank of this matrix is equal to 3. When the data are corrupted by noise we have $\det(\mathbf{H}_{ij}^T - \mathbf{I}) \neq 0$ and an approximate solution must be found. It is well known that in the latter case, the solution of eq. (24) is the eigenvector of $\mathbf{B}^T\mathbf{B}$ associated to its smallest eigenvalue. In practice we consider the singular value decomposition of \mathbf{B} :

$$\mathbf{B} = \mathbf{U}\mathbf{D}(\alpha, \beta, \gamma, \delta)\mathbf{V}^T$$

where \mathbf{D} is a diagonal matrix with $\alpha \geq \beta \geq \gamma \geq \delta \geq 0$ being the singular values of \mathbf{B} . The sought vector $(\mathbf{a}^T \ a)$ is simply the fourth column of the 4×4 orthogonal matrix \mathbf{V} .

The affine calibration method just described is only valid for general rigid motions. Indeed, for pure translations, pure rotations, or planar motions the null space of $\mathbf{D}_{ij} - \mathbf{I}$ is a 2-dimensional space – a pencil of planes.

3.2 Metric calibration

The structure of \mathbf{H}_{PE} given by eq. (22) allows us to write matrix \mathbf{H}_{12} as a function of \mathbf{K} , $(\mathbf{a}^T \ a)$ and \mathbf{D}_{12} . Eq. (1) becomes:

$$\begin{aligned} \mathbf{H}_{12} &= \begin{pmatrix} \overline{\mathbf{H}} & \mathbf{h} \\ \mathbf{k}^T & h \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{K}\mathbf{R}_{12}\mathbf{K}^{-1} + \mathbf{K}\mathbf{t}_{12}\mathbf{a}^T & a\mathbf{K}\mathbf{t}_{12} \\ \frac{1}{a}(-\mathbf{a}^T\mathbf{K}\mathbf{R}_{12}\mathbf{K}^{-1} - \mathbf{a}^T\mathbf{K}\mathbf{t}_{12}\mathbf{a}^T + \mathbf{a}^T) & -\mathbf{a}^T\mathbf{K}\mathbf{t}_{12} + 1 \end{pmatrix} \end{aligned} \quad (25)$$

By simple algebraic manipulations we obtain an expression for the infinite homography between the images of the left camera, before and after the rigid motion (eq. (2)):

$$\mathbf{G}_{12} = \mathbf{K}\mathbf{R}_{12}\mathbf{K}^{-1} = \overline{\mathbf{H}} - \frac{1}{a}\mathbf{h}\mathbf{a}^T \quad (26)$$

3.3 Error analysis

In this section we analyze the relationship between errors associated with affine and metric calibration and errors associated with Euclidean reconstruction. We show that, independently of the calibration method being used, affine calibration has stronger impact than metric calibration. We consider again the relationship between the projective and Euclidean homogeneous coordinates of the same 3-D point, $\mathbf{N} = \mathbf{H}_{PE}\mathbf{M}$ and we write the 4-vector \mathbf{M} as $\mathbf{M}^T = (\overline{\mathbf{M}}^T \ U)$. If $\mathbf{a}^T\overline{\mathbf{M}} + aU \neq 0$ then the point in question is not a point at infinity and we can write $\mathbf{N}^T = (\overline{\mathbf{N}}^T \ 1)$ and we obtain:

$$\overline{\mathbf{N}} = \frac{1}{\mathbf{A}^T \mathbf{M}} \mathbf{K}^{-1} \overline{\mathbf{M}} \quad (27)$$

where $\mathbf{A}^T = (\mathbf{a}^T \ a)$ denotes the fourth row of matrix \mathbf{H}_{PE} , i.e., the plane at infinity.

Let $\hat{\mathbf{H}}_{PE}$ be an estimation of \mathbf{H}_{PE} . We obtain the following estimated Euclidean coordinates:

$$\hat{\overline{\mathbf{N}}} = \frac{1}{\hat{\mathbf{A}}^T \mathbf{M}} \hat{\mathbf{K}}^{-1} \overline{\mathbf{M}} \quad (28)$$

By eliminating $\overline{\mathbf{M}}$ between eqs. (27) and (28) we obtain:

$$\hat{\mathbf{N}} = \frac{\mathbf{A}^T \mathbf{M}}{\hat{\mathbf{A}}^T \mathbf{M}} \hat{\mathbf{K}}^{-1} \mathbf{K} \overline{\mathbf{N}}$$

The matrix \mathbf{K} of intrinsic parameters is the one given by eq. (7). Since the image skew (r) is generally negligible, for the sake of this error analysis we consider a four-parameter camera model:

$$\mathbf{K} = \begin{pmatrix} \alpha & 0 & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

The estimated camera parameters are:

$$\hat{\mathbf{K}} = \begin{pmatrix} \alpha + d\alpha & 0 & u_0 + du_0 \\ 0 & \beta + d\beta & v_0 + dv_0 \\ 0 & 0 & 1 \end{pmatrix}$$

The estimated infinity plane is $\hat{\mathbf{A}} = \mathbf{A} + d\mathbf{A}$. By using first order Taylor expansion we obtain:

$$\hat{\mathbf{N}} = \overline{\mathbf{N}} + \varepsilon_M \overline{\mathbf{N}} + d\mathbf{I} \overline{\mathbf{N}} \quad (29)$$

The matrix $d\mathbf{I}$ and ε_M are given by:

$$d\mathbf{I} = \begin{pmatrix} -\frac{d\alpha}{\alpha} & 0 & -\frac{du_0}{\alpha} \\ 0 & -\frac{d\beta}{\beta} & -\frac{dv_0}{\beta} \\ 0 & 0 & 0 \end{pmatrix}$$

$$\varepsilon_M = \frac{d(\mathbf{A}^T \mathbf{M})}{\mathbf{A}^T \mathbf{M}} = \frac{d(\mathbf{a}^T \overline{\mathbf{M}}) + da U}{\mathbf{a}^T \overline{\mathbf{M}} + a U}$$

Numerically, α and β are one order of magnitude greater than the image center coordinates, u_0 and v_0 . Therefore the entries $\frac{du_0}{\alpha}$ and $\frac{dv_0}{\beta}$ are one order of magnitude smaller than $\frac{d\alpha}{\alpha}$ and $\frac{d\beta}{\beta}$ and can be omitted. Without loss of generality one may assume that $\frac{d\alpha}{\alpha} \approx \frac{d\beta}{\beta} = \varepsilon_f$. The relationship between the “true” Euclidean coordinates and the estimated coordinates becomes:

$$\begin{pmatrix} \hat{x} \\ \hat{y} \\ \hat{z} \end{pmatrix} = \begin{pmatrix} 1 + \varepsilon_M + \varepsilon_f & 0 & 0 \\ 0 & 1 + \varepsilon_M + \varepsilon_f & 0 \\ 0 & 0 & 1 + \varepsilon_M \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (30)$$

Notice that the projective-to-affine error ε_M depends on the accuracy with which the infinity plane is estimated and on the projective coordinates of the reconstructed point.

The affine-to-Euclidean error depends only on the accuracy with which the focal length is estimated. Errors associated with the position of the optical center have a smaller effect onto the Euclidean reconstruction.

4 A linear method for estimating 3-D collineations

The self-calibration and Euclidean reconstruction method described in the previous section relies on proper estimation of collineations between two projective reconstructions, i.e., the matrices \mathbf{H}_{ij} . More generally, let \mathbf{H} be a collineation mapping the points $\mathbf{X}_1, \dots, \mathbf{X}_m$ onto the points $\mathbf{Y}_1, \dots, \mathbf{Y}_m$:

$$\mu_i \mathbf{Y}_i = \mathbf{H} \mathbf{X}_i \quad (31)$$

Given the homogeneous coordinates of \mathbf{X}_i and \mathbf{Y}_i ($X_i^{(j)}$ is the j^{th} coordinate of the 4-vector \mathbf{X}_i), the classical way to estimate the entries of \mathbf{H} is to eliminate the scale factors μ_i . A homogeneous linear system in the entries of \mathbf{H} is thus obtained. This system can be solved when $m \geq 5$ point correspondences are available and with an additional constraint such as $\sum h_{ij} = 1$.

An alternative solution is to estimate simultaneously the entries of \mathbf{H} and the scale factors μ_1, \dots, μ_m . The equation above can be decomposed into four distinct linear constraints and, for example, the first of these linear constraints can be written as:

$$h_{11}X_i^{(1)} + h_{12}X_i^{(2)} + h_{13}X_i^{(3)} + h_{14}X_i^{(4)} + \dots - \mu_i Y_i^{(1)} = 0$$

Without loss of generality we fix one of the scale factors: $\mu_m = 1$. Therefore we have 16 unknowns for the entries of \mathbf{H} and $m - 1$ unknown scale factors. The m equations (31) can be written as a linear system $\mathbf{C}\mathbf{s} = \mathbf{r}$ with $\mathbf{s} = (h_{11} \dots h_{44} \mu_1 \dots \mu_{m-1})^T$, $\mathbf{r} = (\underbrace{0 \dots 0}_{4 \times (m-1)} Y_m^{(1)} Y_m^{(2)} Y_m^{(3)} Y_m^{(4)})^T$, and:

$$\mathbf{C} = \begin{pmatrix} \mathbf{E}_1 & -\mathbf{Y}_1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{E}_2 & \mathbf{0} & -\mathbf{Y}_2 & \dots & \mathbf{0} \\ \vdots & & & & \vdots \\ \mathbf{E}_{m-1} & \mathbf{0} & \mathbf{0} & \dots & -\mathbf{Y}_{m-1} \\ \mathbf{E}_m & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \end{pmatrix}$$

The 4×16 matrices \mathbf{E}_i are defined by:

$$\mathbf{E}_i = \begin{pmatrix} \mathbf{X}_i^T & \mathbf{0}^T & \mathbf{0}^T & \mathbf{0}^T \\ \mathbf{0}^T & \mathbf{X}_i^T & \mathbf{0}^T & \mathbf{0}^T \\ \mathbf{0}^T & \mathbf{0}^T & \mathbf{X}_i^T & \mathbf{0}^T \\ \mathbf{0}^T & \mathbf{0}^T & \mathbf{0}^T & \mathbf{X}_i^T \end{pmatrix}$$

This linear system consists of $4m$ equations. Since there are $16+(m-1) = 15+m$ unknowns, we must have $m \geq 5$. This linear system can be solved using standard techniques provided that the 3-D points are not coplanar.

One way to assess the quality of the estimated collineation $\hat{\mathbf{H}}$ is to compare the projections of $\hat{\mathbf{Y}}_i = \hat{\mathbf{H}}\mathbf{X}_i$ and of $\hat{\mathbf{X}}_i = \hat{\mathbf{H}}^{-1}\mathbf{Y}_i$ with the true image points. Let \mathbf{x}_i and \mathbf{x}'_i be the true image points (in the left and right images) from which the 3-D point \mathbf{X}_i was reconstructed, and let $\hat{\mathbf{x}}_i$ and $\hat{\mathbf{x}}'_i$ be the projections of $\hat{\mathbf{X}}_i$. Similarly we define \mathbf{y}_i , \mathbf{y}'_i , $\hat{\mathbf{y}}_i$ and $\hat{\mathbf{y}}'_i$:

$$\begin{aligned} \hat{\mathbf{x}}_i &= \mathbf{P}\hat{\mathbf{H}}^{-1}\mathbf{Y}_i \\ \hat{\mathbf{x}}'_i &= \mathbf{P}'\hat{\mathbf{H}}^{-1}\mathbf{Y}_i \\ \hat{\mathbf{y}}_i &= \mathbf{P}\hat{\mathbf{H}}\mathbf{X}_i \\ \hat{\mathbf{y}}'_i &= \mathbf{P}'\hat{\mathbf{H}}\mathbf{X}_i \end{aligned}$$

With the notation $\mathbf{y}^T = (\bar{\mathbf{y}}^T \ 1)$ let $d(\bar{\mathbf{y}}, \hat{\bar{\mathbf{y}}})$ denote the Euclidean distance between the image points \mathbf{y} and $\hat{\mathbf{y}}$. The quality of the collineation is assessed by the following quadratic error function:

$$f(\hat{\mathbf{H}}, \hat{\mathbf{H}}^{-1}) = \frac{1}{4m} \sum_{i=1}^m \left(d(\bar{\mathbf{x}}_i, \hat{\bar{\mathbf{x}}}_i)^2 + d(\bar{\mathbf{x}}'_i, \hat{\bar{\mathbf{x}}}'_i)^2 + d(\bar{\mathbf{y}}_i, \hat{\bar{\mathbf{y}}}_i)^2 + d(\bar{\mathbf{y}}'_i, \hat{\bar{\mathbf{y}}}'_i)^2 \right) \quad (32)$$

Finally the error function defined by eq. (32) can be used to estimate the collineation by minimizing the following non-linear criteria:

$$\min_{\mathbf{H}, \mathbf{H}'} \left(f(\mathbf{H}, \mathbf{H}') - \eta \|\mathbf{H}\mathbf{H}' - \mathbf{I}\|^2 \right) \quad (33)$$

where $\|\mathbf{H}\mathbf{H}' - \mathbf{I}\|^2$ is a penalty function and η is a real positive number. A high numerical value for η guarantees that $\mathbf{H}' = \mathbf{H}^{-1}$.

We implemented both the linear and non-linear methods for estimating \mathbf{H} and carried out a large number of experiments in order to compare the quality of the results. Figure 1

plots the errors given by eq. (32) as a function of image noise for both the linear (L) and non-linear (N) resolution methods. This noise has Gaussian distribution with standard deviation varying from 0 to 2 pixels.

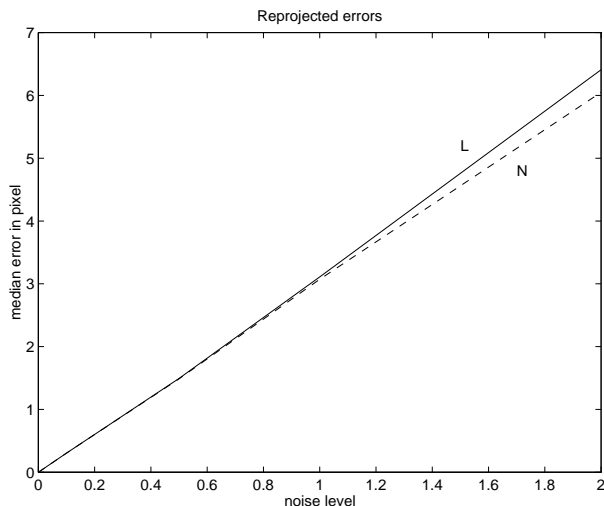


Figure 1: The quality of the collineation linearly degrades in the presence of image noise. The behavior of the linear (L) and non-linear (N) methods are almost identical. This plot allows us to estimate, *a posteriori*, the level of noise which is present in the data.

5 Implementation

The method described above was implemented and tested with both simulated and experimental data. Simulated data allows us to systematically study the sensitivity of the method with respect to image noise and to assess the conditions under which reliable results may be expected.

We used two types of experimental data: “calibrated data” and natural data. Calibrated data consists of images of a 3-D calibrated object. Since the metrics of this object is perfectly known, we can use standard camera calibration techniques and compare the results obtained with our self-calibration procedure with standard off-line camera calibration methods. Moreover, the calibrated data is so accurate that the camera parameters obtained with these data and with a classical camera calibration method may well be viewed as the ground-truth.

The self-calibration and reconstruction method can be summarized in the following steps:

1. Move the stereo rig and for each position perform a projective reconstruction;
2. For each motion between position i and position $i+1$ compute the collineation $\mathbf{H}_{i\ i+1}$;
3. Solve eq. (24) to find the plane at infinity;
4. For each motion use eq. (26) to compute the infinite homography associated with the left camera;
5. Combine all the motions to estimate the intrinsic camera parameters using eq. (3);
6. Multiply matrix \mathbf{H}_{PE} thus obtained with the projective coordinates of a 3-D point to get its Euclidean coordinates.

5.1 Noise sensitivity analysis

The simulated data consists of 41 3-D points. These points are projected onto the images of a stereo rig with known epipolar geometry and known intrinsic camera parameters. Sets of 2-D points are obtained by simulating various rigid motions.

The noise added to the image points is Gaussian with standard deviation varying from 0.05 to 1 pixel. Each experiment consists of 100 trials at some fixed standard deviation and the median error over these trials is computed.

We studied the behavior of the method as a function of image noise and as a function of the number of motions of the stereo rig. According to section 2.1 three camera models are studied: a camera with three parameters (P3), a camera with four parameters (P4) and a camera with five parameters (P5). Notice that in the case of a single motion, only the P3 and P4 camera models can be used. The simulated stereo rig has the following intrinsic parameters (left camera):

$$\mathbf{K} = \begin{pmatrix} 715 & 0 & 240 \\ 0 & 995 & 275 \\ 0 & 0 & 1 \end{pmatrix}$$

Figure 2 shows the relative median error associated with the intrinsic camera parameters for 3 motions of the stereo rig and for the 3 possible camera models. Figure 3 shows the

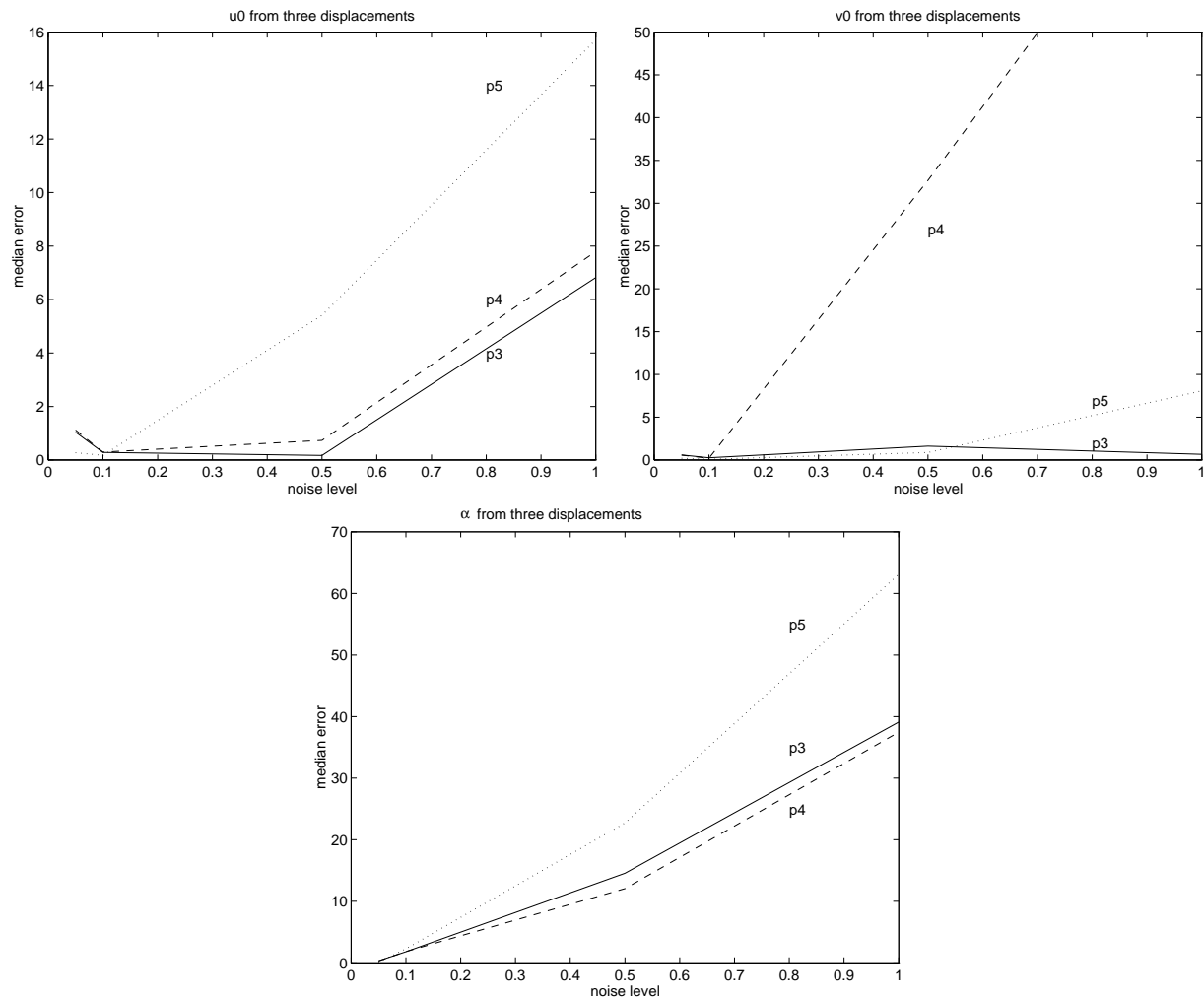


Figure 2: The median errors associated with the estimation of the intrinsic camera parameters as a function of image noise. Three motions were simulated for the scope of these experiments.

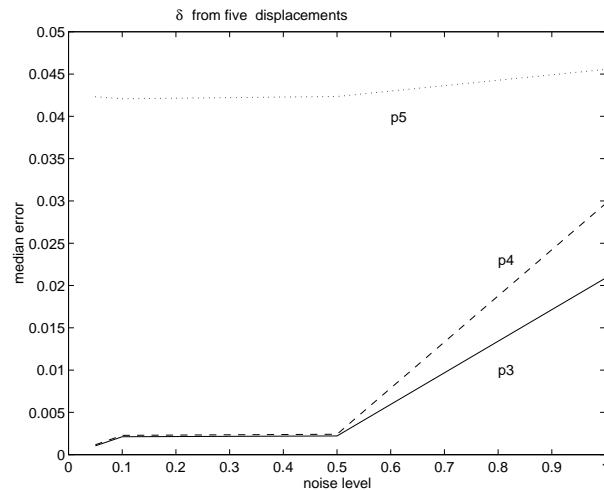


Figure 3: The relative error in reconstruction (median error) as a function of image noise. This experiment shows that Euclidean reconstruction tolerates “bad” camera calibration.

discrepancy between the true 3-D Euclidean points and the estimated ones for 5 motions of a stereo rig and for the 3 possible camera models.

In order to have a more global view we plotted the values obtained for camera calibration over all the trials: both the camera model (P3, P4, and P5) and the number of motions vary but, for each plot, the standard deviation of the added noise is fixed. The distributions that we obtained for the camera parameters are shown on Figure 4 and Figure 5.

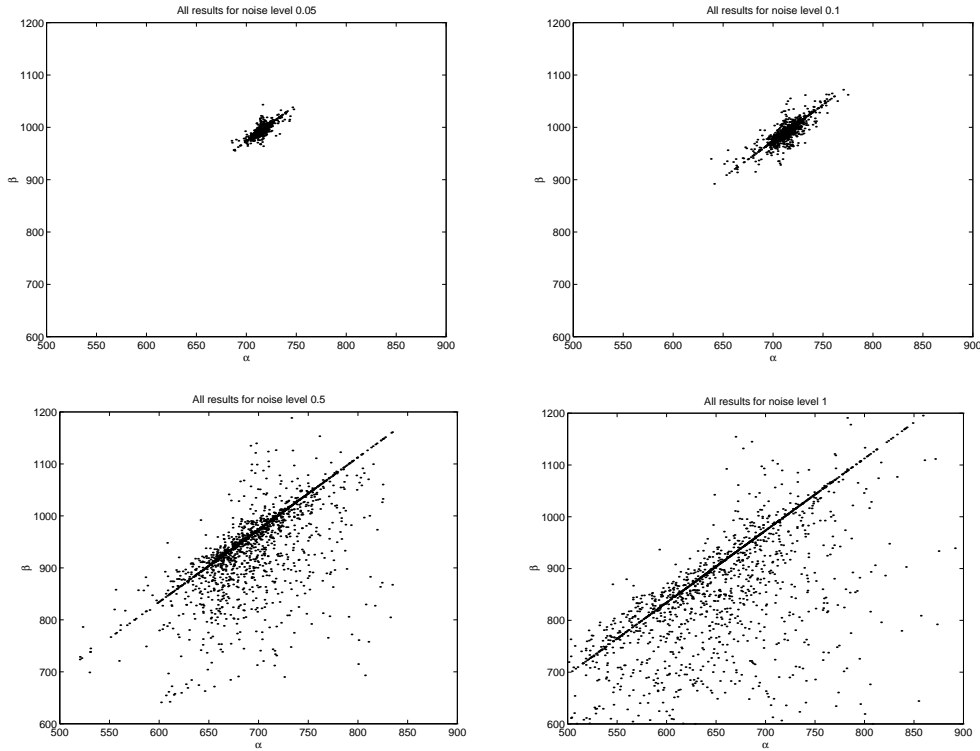


Figure 4: The statistical distribution of the horizontal and vertical image scale factors computed over a large number of experiments and for 4 levels of image noise: 0.05, 0.1, 0.5, and 1 pixel.

5.2 Experiments with real data data

As already mentioned, we tested our method over two types of real data: “calibrated data” and natural data. Calibrated data consists of a set of 100 circular targets evenly distributed over the three planes of a calibrated object. The images of these targets are detected and their centers are localized with an accuracy of 0.05 pixels. These data are called calibrated because the 3-D positions of the targets’ centers are known in an object-oriented Euclidean frame with an accuracy of 0.02mm.

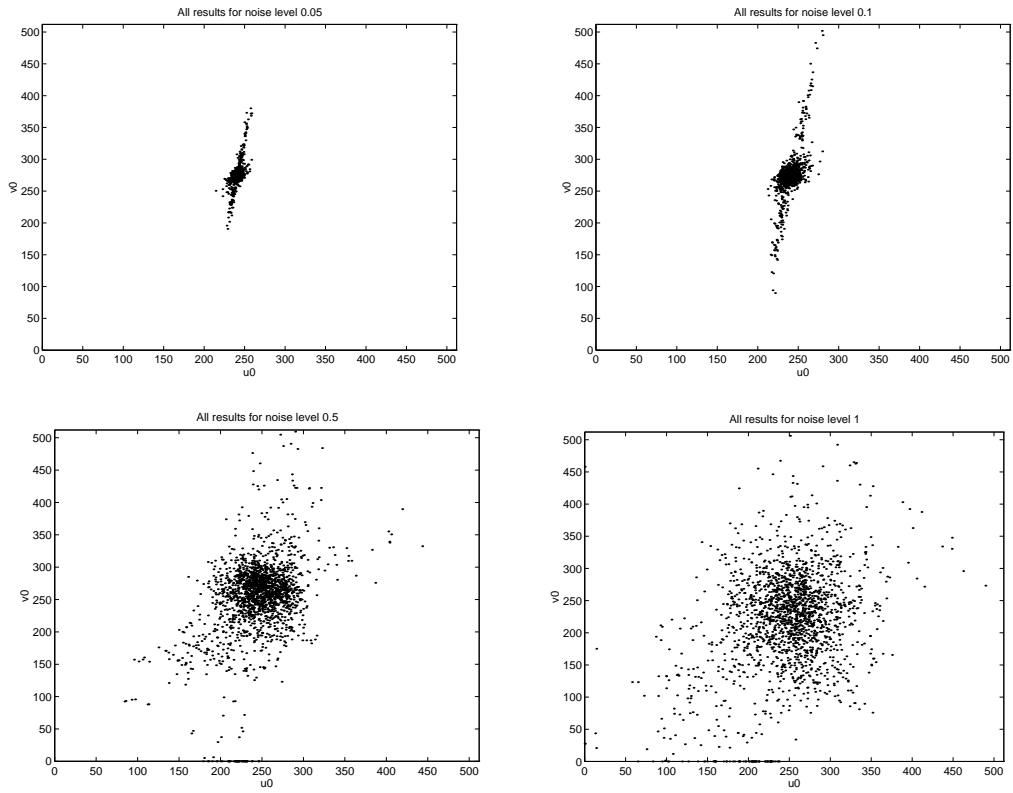


Figure 5: Same as the previous figure but for the image center.

We gathered three image pairs of this object, Figure 6. First we calibrated the left camera using 2-D/3-D point matches. The result of this classical camera calibration procedure is displayed on Table 1, first row. Because of the accuracy of this calibration, these parameters are considered as the ground-truth. Second, we applied our calibration method to three image pairs and their corresponding 2-D/2-D matches and obtained calibration results which are displayed on Table 1, rows 2, 3, and 4.

Finally, we gathered four image pairs of a natural scene, the second pair being displayed on Figure 7. Points are detected and localized using a standard point-of-interest operator. These points are matched, between the left and right images for each image pair and between consecutive image pairs. There are approximately 90 matched points available to compute a collineation between two image pairs. These collineations ($\mathbf{H}_{12}, \mathbf{H}_{23}, \mathbf{H}_{34}$) are estimated using the linear method described in section 4. The median error associated with the estimation of these matrices is of approximately 1.26 pixels. The median-error-versus-image-noise curve of Figure 1 allows us to estimate the level of noise associated with these “natural” data – in this case the standard deviation of the noise is 0.5 pixels. The calibration results obtained with these data are shown on Table 1, rows 5 to 7.

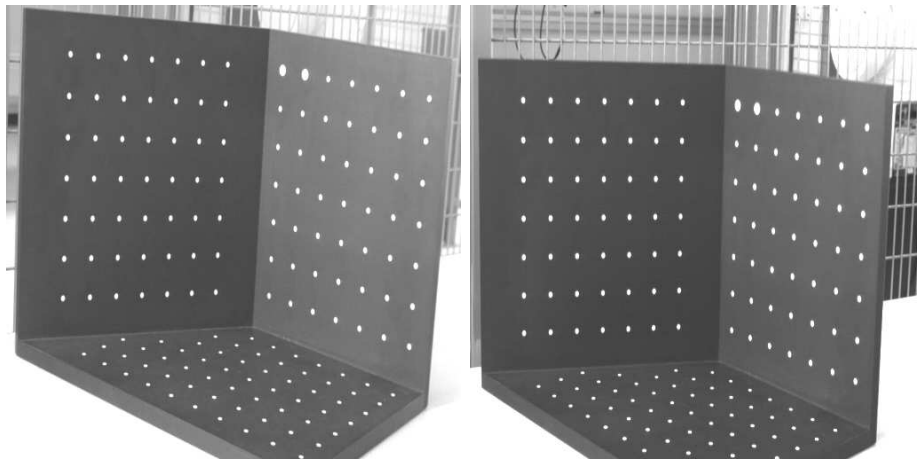


Figure 6: An image pair of the calibrated object used for both off-line calibration and self-calibration

6 Discussion

We described a method for recovering camera calibration and metric reconstruction from general rigid motions of an uncalibrated stereo rig. We showed how to reliably extract the plane at infinity from several rigid motions and how to convert the affine calibration

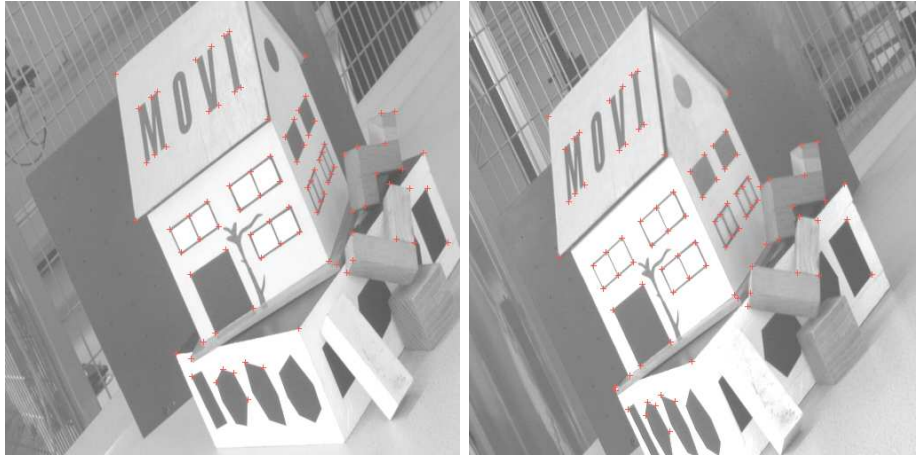


Figure 7: One among the four image pairs of a natural scene used in our experiments.

Method	α	$k\alpha$	u_0	v_0	Comments
Off-line calibration	1045	1540	252	245	Known 3-D geometry
P3	1042	1531	253	240	Two motions and calibrated data
P4	1042	1527	250	238	
P5	1035	1522	250	231	
P3	928	1364	270	122	Three motions and natural data
P4	877	1368	246	120	
P5	969	1426	247	208	

Table 1: Results of off-line calibration (first row), self-calibration with “calibrated” data (rows 2 to 4) and with “natural” data (rows 5 to 7).

thus obtained into metric calibration. An error analysis emphasizes the importance of affine calibration over metric calibration. We proposed a linear method for computing the collineation between two projective reconstruction and we showed that this linear method performs almost as well as a non-linear minimization method. The quality of this linear estimation degrades linearly with noise and allows to determine, a posteriori, the amount of noise associated with image features. Extensive experiments with noisy simulated data allowed a statistical characterization of the behavior of the method and a noise sensitivity analysis. Based on this one can conclude that the method tolerated Gaussian noise with standard deviation up to 0.5 pixels.

The experiments performed with calibrated data allowed both a comparison with off-line camera calibration techniques and a validation of the statistical behavior in the presence of noise. Indeed, the calibrated data have an accuracy of 0.05 pixels. It can be seen from Figures 4 and 5 that at this level of noise the results are quite reliable and that the intrinsic parameters thus derived may well be considered as the ground truth.

The experiments performed with natural data confirmed as well the error analysis, the statistical behavior, and the noise sensitivity analysis.

The method has been extensively evaluated with three camera models. Indeed, the question of whether one should use a 5-parameter, 4-parameter, or 3-parameter camera was somehow open. The statistical analysis does not reveal that one model is more resistant to noise than another. In practice we believe that a 4-parameter camera is the most suited model.

References

- [1] M. Armstrong, A. Zisserman, and R. Hartley. Self-calibration from image triplets. In Buxton-Cipolla, editor, *Computer Vision – ECCV’96, Proceedings Fourth European Conference on Computer Vision, Cambridge, England*, pages 3–16. Springer Verlag, April 1996.
- [2] 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.

- [3] F. Devernay and O. Faugeras. From projective to euclidean reconstruction. In *Proceedings Computer Vision and Pattern Recognition Conference*, San Francisco, CA., 1996.
- [4] O. D. Faugeras, Q. T. Luong, and S. J. Maybank. Camera self-calibration: Theory and experiments. In G. Sandini, editor, *Computer Vision – ECCV 92, Proceedings Second European Conference on Computer Vision, Santa Margherita Ligure, May 1992*, pages 321–334. Springer Verlag, May 1992.
- [5] O.D. Faugeras. Stratification of 3-d vision: projective, affine, and metric representations. *Journal of the Optical Society of America A*, 12(3):465–484, March 1995.
- [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] R. I. Hartley. Self-calibration from multiple views with a rotating camera. In *Proc. Third European Conference on Computer Vision*, pages 471–478, Stockholm, Sweden, May 1994.
- [8] R. I. Hartley. In defence of the 8-point algorithm. In *Proceedings Fifth International Conference on Computer Vision*, pages 1064–1070, Cambridge, Mass., June 1995. IEEE Computer Society Press, Los Alamitos, Ca.
- [9] R. I. Hartley. Kruppa’s equations derived from the fundamental matrix. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(2):133–136, February 1997.
- [10] Q-T. Luong. *Matrice Fondamentale et Autocalibration en vision par ordinateur*. PhD thesis, Université de Paris Sud, Orsay, December 1992.
- [11] Q-T. Luong and O. D. Faugeras. The fundamental matrix: Theory, algorithms, and stability analysis. *International Journal of Computer Vision*, 17(1):43–75, 1996.
- [12] Q-T. Luong and T. Viéville. Canonic representations for the geometries of multiple projective views. In Jan-Olof Eklundh, editor, *Computer Vision – ECCV 94, Proceedings Third European Conference on Computer Vision*, volume 1, pages 589–599. Springer Verlag, Stockholm, Sweden, May 1994.

- [13] S. J. Maybank and O. D. Faugeras. A theory of self calibration of a moving camera. *International Journal of Computer Vision*, 8(2):123–151, August 1992.
- [14] 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.
- [15] Z. Zhang, R. Deriche, O. D. Faugeras, and Q-T. Luong. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence*, 78(1–2):87–119, October 1995.
- [16] 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.