

Title: **Calibrating a video camera pair with a rigid bar**

Authors: N. Alberto Borghese⁺ and P. Cerveri^{*+}

Affiliations: ⁺Laboratory of Human Motion Study and Virtual Reality,
Istituto Neuroscienze e Bioimmagini CNR, Via f.lli Cervi, 93 -
20090 Segrate (Milano), I – borghese@inb.mi.cnr.it.

^{*}Department of Bioengineering, Centro di Bioingegneria,
Fondazione ProJuventute, Politecnico di Milano, Pzza Leonardo
da Vinci, 32, 20100 Milano, I – cerveri@biomed.polimi.it.

Running Headline: Bar calibration #2

Acknowledgment: This work has been mainly carried out at the Laboratory of
Human Motion Study and Virtual Reality, INB-CNR.

Mailing address:

Borghese Nunzio Alberto Dr. Eng.
Laboratory of Human Motion Study and Virtual Reality,
Istituto Neuroscienze e Bioimmagini CNR,
LITA - Via F.lli Cervi, 93
20090 Segrate - Milano - I
borghese@inb.mi.cnr.it

Number of pages: 26

Number of tables: 2

Number of figures: 5

Abstract

In this paper a new procedure to determine all the geometrical parameters of a stereo-system is presented. It is based on surveying a rigid bar carrying two markers on its extremities moved inside the working volume and it does not require grids or complex calibration structures. The external parameters are estimated through the epipolar geometry up to a scale factor which is determined from the true length of the bar. The focal lengths are determined using the properties of the absolute conic in the projective space. The principal points are computed through a non-linear minimization carried out through an evolutionary optimization. The accuracy of the method is assessed on real data and it compares favorably with that obtained through classical approaches based on control points of known 3D coordinates.

Keywords: Epipolar geometry, Evolution strategy, Calibration, Fundamental matrix.

1. Introduction

Classical calibration techniques are based on surveying a 3D distribution of control points of known position^(1,2,3). To achieve a high accuracy, the control points should be positioned with extreme precision and distributed over the entire working volume⁽⁴⁾; therefore large, difficult to move and expensive calibration structures are usually required.

A different approach is represented by self-calibration procedures which do not require to know in advance the 3D position of the control points. Classical photogrammetry offers the most general solution through the method of bundles adjustment^(5,6). This allows to determine the 3D coordinates of the control points and the internal and external parameters of the set-up at the same time through a non-linear minimization of the distance between the reprojected and the measured 2D positions of the control points. In this method, the dimension of the design matrix and the number of unknowns increases linearly with the number of the control points leading to a huge computational load. Moreover, singularities in the design matrix can easily occur, and these may hamper the accuracy in the parameters. To avoid this, the distribution of the 3D control points should be carefully examined before starting the estimate. Finally, bundle adjustment requires a good initialisation of both the parameters and the 3D coordinates of the control points.

A different approach to self-calibration has been developed under the framework of the so called "Structure from Motion" (SfM) problem (3,7,8,9,10). In this framework, the scene is surveyed by a moving camera; its motion as well as the 3D position of a set of points in the scene

are determined from the image sequence. The internal parameters are supposed known. A linear solution to SfM problems was proposed by Longuet-Higgings⁽¹¹⁾: it is based on the epipolar constraint which expresses the condition that the two straight lines, through the 2D measured positions of a point and the perspective centers, lie on the same plane: the epipolar plane (cf. Fig. 1). This solution allows to keep the design matrix small (9x9). SfM solutions have been recently extended by Hartley^(12,13) and Faugeras and co-workers^(14,15,16) to estimate also the internal parameters, provided that the sequence is constituted of a sufficient number of images.

The SfM framework has been linked to stereo-calibration by Borghese and Perona⁽¹⁷⁾ who made the observation that the two cameras of a stereo-pair can be associated to a sequence of two images, taken by the moving camera from two different positions. Under this hypothesis, only the focal lengths can be determined by SfM solutions^(3, 12) and the position of the two principal points has to be determined with a different procedure. As a first approximation, they can be assumed coincident with the image center, but, due to imperfect assembling of the optical system, they can be offset of several pixels. When high noise levels in the measurement is present, the contribution of the mislocation of the principal point to the 3D reconstruction can be neglected⁽¹⁸⁾ and accuracy sufficient to certain applications like human computer interface can be attained. However, this approximation is not adequate when high accuracy is required. Several techniques have been proposed to determine the true position of the principal point. In Wei and Ma⁽¹⁹⁾ the cross-ratio invariance over projection is exploited but to get a reliable estimate of the principal point, pure radial distortions on the lens is required, which is not often the case. Different solutions are based on ad-hoc devices like a rigid cross filmed on the field⁽⁸⁾ or laser based systems^(20,21). The use of additional devices make this kind of calibration procedure less friendly and less easy to be performed.

In this paper a novel method to estimate all the calibration parameters of a stereo pair. It does not require any initialisation (differently from bundle adjustment) and it allows to determine the internal and external parameters included the two principal points (differently from SfM approaches), using control points whose 3D position has not to be measured. The results of real calibrations carried out on the field are reported and discussed.

2. Geometrical Background

2.1 Geometrical model of the stereo system

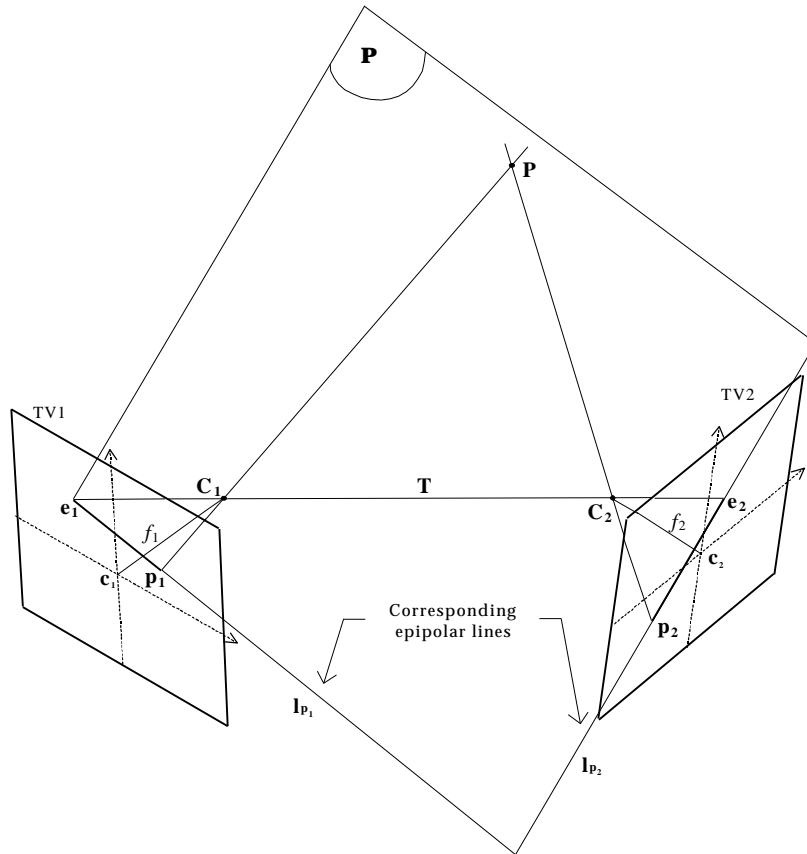


Figure 1. The geometrical arrangement of a stereo-pair. The point \mathbf{P} is projected in \mathbf{p}_1 and \mathbf{p}_2 on the image plane of the two cameras. The optical axis intersects the image plane at a point $\mathbf{c}_j(x_{o_j}, y_{o_j})$ called principal point. The distance between the perspective center and the image plane, f_j , is the focal length of the camera. The epipolar plane, Π , contains the points \mathbf{P} , \mathbf{p}_1 , \mathbf{p}_2 , and the line $\overline{C_1 C_2}$ (the relative position \mathbf{T}), which joins the perspective centres of the two cameras. The points of intersection between $\overline{C_1 C_2}$ and the image planes are the epipoles, \mathbf{e}_1 and \mathbf{e}_2 . The intersection between Π and the two image planes originates two lines, \mathbf{l}_{p_1} and \mathbf{l}_{p_2} , called epipolar lines. They are corresponding epipolar lines in the homography generated by \mathbf{P} , \mathbf{C}_1 and \mathbf{C}_2 .

The projection of a 3D point, $\mathbf{P}(X,Y,Z)$, on the image plane of the video camera j , $\mathbf{p}_j(x,y)$, is described by the perspective equations, which, in homogeneous notation, are⁽³⁾:

$$\mathbf{p} = \mathbf{K M D P} \quad \text{with} \quad \mathbf{B} = \mathbf{K M} \tag{1}$$

where:

$$\mathbf{K} = \begin{bmatrix} -f & 0 & x_o \\ 0 & -f & y_o \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

f is the focal length of the camera and (x_o, y_o) are the coordinates of the principle point, $\mathbf{c}(x_o, y_o)$ and f constitute the internal parameters of the camera (cf. Fig. 1).

$$\mathbf{M} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (3)$$

$$\mathbf{D} = \begin{bmatrix} \mathbf{R} & -\mathbf{RT} \\ \mathbf{0} & 1 \end{bmatrix} \quad (4)$$

\mathbf{R} and \mathbf{T} are the orientation (3x3 matrix) and the location vector of the camera with respect to a certain reference frame: they constitute the external parameters. Calibration is here defined as the determination of the internal and external parameters for the stereo-pair.

Whenever the two focal length and the two principle point are known, the normalised target coordinates of a point can be introduced. They are:

$$\hat{\mathbf{p}}_j = \begin{bmatrix} \frac{x_j - x_{o_j}}{-f_j} \\ \frac{y_j - y_{o_j}}{-f_j} \\ 1 \end{bmatrix} \quad (5)$$

and they will be used in Section 2.4 to estimate the external parameters. Whenever only the principle point is known, the offset coordinates are defined as:

$$\tilde{\mathbf{p}}_j = \begin{bmatrix} x_j - x_{o_j} \\ y_j - y_{o_j} \\ 1 \end{bmatrix} \quad (6)$$

They will be adopted in Section 2.3 to estimate the two focal lengths.

For the sake of convenience the absolute reference frame is taken solid with one of the two cameras with the X,Y axes parallel to those of its image plane. With this choice, equation (1) for the two cameras becomes:

$$\mathbf{p}_1 = \mathbf{K}_1 \mathbf{M} \mathbf{P} \quad (7a)$$

$$\mathbf{p}_2 = \mathbf{K}_2 \mathbf{M} \mathbf{D} \mathbf{P} \quad (7b)$$

These show that the internal and the external parameters can be factorised into the product of a matrix containing the internal parameters, \mathbf{K} , by a matrix containing of the external ones, \mathbf{D} .

2.2 The epipolar geometry

The epipolar geometry is based on the observation that each 3D point, \mathbf{P} , and its projections, \mathbf{p}_1 and \mathbf{p}_2 , lie on the same plane Π , called epipolar plane (cf. Fig. 1). This plane intersects the two image planes along two lines, \mathbf{l}_{p1} and \mathbf{l}_{p2} , called epipolar lines. For construction, a point \mathbf{p}_1 , belonging to \mathbf{l}_{p1} , will correspond to \mathbf{l}_{p2} ; and a point \mathbf{p}_2 , belonging to \mathbf{l}_{p2} , will correspond to \mathbf{l}_{p1} . This condition is expressed analytically as follows:

$$\mathbf{l}_{p2} = \mathbf{F}\mathbf{p}_1 \quad (8a)$$

$$\mathbf{l}_{p1} = \mathbf{F}^T\mathbf{p}_2 \quad (8b)$$

where \mathbf{F} is the Fundamental matrix⁽³⁾. As the scalar product between a point and a line to which it belongs to, is equal to 0, the following relationship holds:

$$\mathbf{p}_2^T \mathbf{l}_{p2} = 0 \quad (9)$$

and from equations (8):

$$\mathbf{p}_2^T \mathbf{F}\mathbf{p}_1 = 0 \quad (10)$$

which represents the epipolar plane, Π , identified by \mathbf{p}_1 , \mathbf{p}_2 and $\overline{C_1C_2}$. The epipoles, \mathbf{e}_1 and \mathbf{e}_2 , are the points of intersection of the line through $\overline{C_1C_2}$ and the two image planes. For these points, the epipolar lines, \mathbf{l}_{e1} and \mathbf{l}_{e2} , are reduced to the epipoles themselves; it holds:

$$\mathbf{F}\mathbf{e}_1 = \mathbf{F}^T\mathbf{e}_2 = 0 \quad (11)$$

This equation will be used to determine the focal lengths in Section 2.3. Matrix \mathbf{F} can also be obtained through different considerations. Imposing that the three vectors $\overline{C_1C_2}$, $\overline{PC_1}$ and $\overline{PC_2}$ are coplanar, the following relationship holds:

$$\overline{PC_2} \times \mathbf{T} \wedge \overline{PC_1} = 0 \quad (12)$$

Defining a skew symmetric matrix \mathbf{S} which represents the elements of \mathbf{T} as:

$$\mathbf{S} = (\mathbf{T} \wedge) \Rightarrow \mathbf{S} = \begin{bmatrix} 0 & -T_z & T_y \\ T_z & 0 & -T_x \\ -T_y & T_x & 0 \end{bmatrix} \quad (13)$$

and taking into account equations (7), the following homogeneous equation is obtained^(9,11):

$$\mathbf{p}_2^T \mathbf{K}_2^{-T} \mathbf{R} \mathbf{S} \mathbf{K}_1^{-1} \mathbf{p}_1 = 0 \quad (14)$$

with:

$$\mathbf{F} = \mathbf{K}_2^{-T} \mathbf{R} \mathbf{S} \mathbf{K}_1^{-1} \quad (15)$$

This contains all the calibration parameters factorised into the product of the internal (matrices \mathbf{K}_1 and \mathbf{K}_2) by the external ones (matrices \mathbf{R} and \mathbf{S}). When the internal parameters are known, the normalized coordinates (equations (2)) can be used, and equation (14) assumes the following shape:

$$\hat{\mathbf{p}}_2^T \mathbf{E} \hat{\mathbf{p}}_1 = 0 \quad \text{with} \quad \mathbf{E} = \mathbf{R} \mathbf{S} \quad (16)$$

\mathbf{E} is named the essential matrix and it contains only the external parameters which can be estimated according to the procedure described in Section 2.4. Algebraic considerations which limit the number of independent parameters which can be estimated from equation (14) to seven are reported in Appendix A. Of these seven parameters, five parameters will be the relative orientation and location up to a scale factor, and the other two will be used to compute the focal lengths⁽²²⁾. No other parameter can be directly estimated with this approach.

2.3 The Fundamental matrix and the focal lengths.

To estimate the internal parameters, the matrices \mathbf{K}_1 and \mathbf{K}_2 in equation (14) are isolated from \mathbf{R} and \mathbf{T} through considerations based on projective geometry. In particular, the correspondence between the lines tangent to the absolute conic is exploited^(12,16). The absolute conic, Σ_∞ , is given by the following equation:

$$\mathbf{P}^T \mathbf{P} = 0 \quad (17)$$

which has no real points (only imagery points). It has the remarkable property to be invariant for Euclidean transformations, and, in particular, for rotation and translation. Σ_∞ is projected onto the image plane of each camera into the conic, σ_j , (cf. Fig. 2) which is described by:

$$\mathbf{p}_j^T \mathbf{K}_j^{-T} \mathbf{K}_j^{-1} \mathbf{p}_j = 0 \tag{18}$$

It can be shown that the problem of determining the internal parameters of a camera is equivalent to finding $\sigma_j^{(23)}$. Let us specialise the matrices \mathbf{K}_1 and \mathbf{K}_2 for the case when the principal points are known and only the focal lengths have to be determined. In this case offset coordinates can be adopted (equations (3)), and \mathbf{K}_j assumes the following shape:

$$\mathbf{K}_j = \begin{bmatrix} -f_j & 0 & 0 \\ 0 & -f_j & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{19}$$

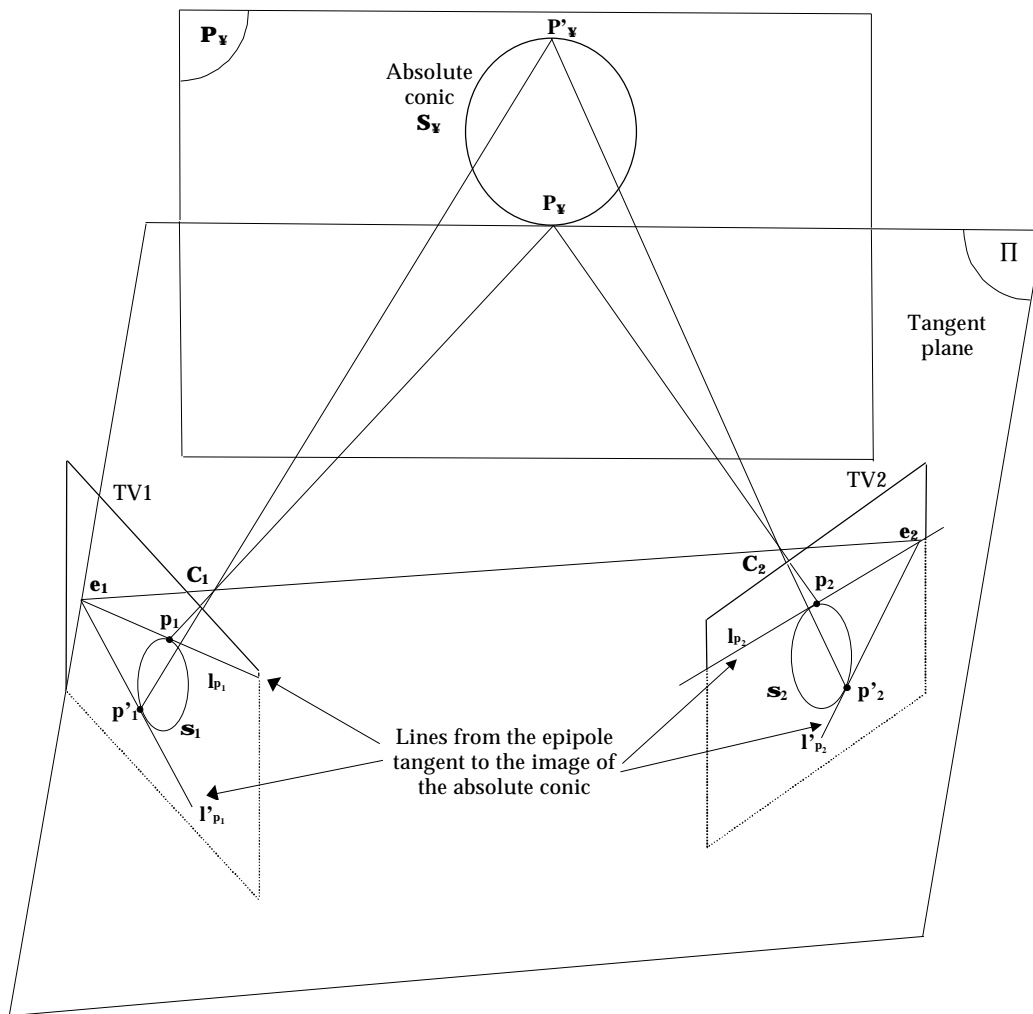


Figure 2. The absolute conic, Σ_∞ , and its projection over the two image planes, σ_1 and σ_2 . The intersection of the two tangent planes with the absolute conic gives the two points \mathbf{P}_Ψ and \mathbf{P}'_Ψ which project into \mathbf{p}_1 and \mathbf{p}'_1 on the first image plane and into \mathbf{p}_2 and \mathbf{p}'_2 on the second one. Through these points the two pairs of epipolar lines, \mathbf{l}_{p_1} and \mathbf{l}'_{p_1} , and \mathbf{l}_{p_2} and \mathbf{l}'_{p_2} are identified.

Equation (18) becomes:

$$\tilde{\mathbf{p}}_j^T \text{diag}(1, 1, f_j^2) \tilde{\mathbf{p}}_j = 0 \quad (20)$$

This matrix represents the imaginary circle, \mathbf{S}_j , centred in the origin which, in homogeneous coordinates is:

$$\sigma_j: \tilde{u}_j^2 + \tilde{v}_j^2 + \tilde{w}_j^2 f_j^2 = 0 \quad (21)$$

As the projection of the absolute conic depends only on the internal parameters (equation (18)), the reference system of each camera can be arbitrarily positioned. In particular, to get a simplified shape for the matrix \mathbf{F} , the reference system is positioned such that the epipole has coordinates $[1, 0, 1]$:

$$\mathbf{e}_1 = \mathbf{e}_2 = [1, 0, 1]^T \quad (22)$$

With this reference system, taking into account equation (11), matrix \mathbf{F} assumes the following simplified shape:

$$\mathbf{F} = \begin{bmatrix} a & b & -a \\ c & d & -c \\ -a & -b & a \end{bmatrix} \quad (23)$$

To compute f_1 and f_2 , the properties of the correspondence between points and epipolar lines described by equations (8) is exploited. Let us consider the lines $\mathbf{l}_{\mathbf{p}_1}$ and $\mathbf{l}'_{\mathbf{p}_1}$ through \mathbf{e}_1 which are tangent to \mathbf{S}_1 (cf. Fig. 2). The two points of tangency are \mathbf{p}_1 and \mathbf{p}'_1 . They identify the polar* of \mathbf{e}_1 with respect to \mathbf{S}_1 which has the following equation:

$$\frac{\mathcal{I}[\mathbf{s}_1]}{\mathcal{I}[u_1]_{e_1}} u_1 + \frac{\mathcal{I}[\mathbf{s}_1]}{\mathcal{I}[v_1]_{e_1}} v_1 + \frac{\mathcal{I}[\mathbf{s}_1]}{\mathcal{I}[w_1]_{e_1}} w_1 = 0 \quad \Rightarrow \quad u_1 + w_1 f_1^2 = 0 \quad (24)$$

The intersection of the line in equation (24) with the conic in equation (21), the coordinates originates two imagery points, \mathbf{p}_1 and \mathbf{p}'_1 , which are computed as:

* The polar of a point, \mathbf{p} , with respect to a conic σ , is the line joining the two points of tangency with σ of the lines through \mathbf{p} .

$$\begin{cases} u_1 + w_1 f_1^2 = 0 \\ u_1^2 + v_1^2 + w_1^2 f_1^2 = 0 \end{cases} \Rightarrow \begin{aligned} u_1 &= -w_1 f_1^2 \\ v_1^2 &= -w_1^2 f_1^2 (1 + f_1^2) \\ w_1 &= w_1 \end{aligned} \quad (25)$$

Introducing the imaginary unit i , and dividing the three homogenous equations by w_1 , the coordinates of \mathbf{p}_1 and \mathbf{p}'_1 can be written as:

$$\mathbf{p}_1 = [-f_1^2, \quad i f_1(f_1^2+1)^{1/2}, \quad 1]^T \quad (26a)$$

$$\mathbf{p}'_1 = [-f_1^2, \quad -i f_1(f_1^2+1)^{1/2}, \quad 1]^T \quad (26b)$$

The points \mathbf{p}_1 and \mathbf{p}'_1 can be transformed through \mathbf{F} into their corresponding epipolar lines, \mathbf{l}_{p_2} and \mathbf{l}'_{p_2} (equations (8)), which will be tangent to σ_2 . Using relations (23) and (26), and taking into account that equations (8) are defined up to a scale factor, equation (8a) can be written for the points \mathbf{p}_1 and \mathbf{p}'_1 as:

$$\mathbf{F}\mathbf{p}_1 = \begin{bmatrix} -1 \\ \frac{-c\Delta + d i f_1}{a\Delta - b i f_1} \\ 1 \end{bmatrix} = \mathbf{l}_2 \quad (27a)$$

$$\mathbf{F}\mathbf{p}'_1 = \begin{bmatrix} -1 \\ \frac{-c\Delta - d i f_1}{a\Delta - b i f_1} \\ 1 \end{bmatrix} = \mathbf{l}'_2 \quad (27b)$$

where $\Delta = (f_1^2+1)^{1/2}$. In compact notation, equations (27) can be written as:

$$\mathbf{F}\mathbf{p}_1 = [-1 \ (\mathbf{m} + i\mathbf{u}) \ 1]^T = \mathbf{l}_2 \quad (28a)$$

$$\mathbf{F}\mathbf{p}'_1 = [-1 \ (\mathbf{m} - i\mathbf{u}) \ 1]^T = \mathbf{l}'_2 \quad (28b)$$

where:

$$\mathbf{m} + i\mathbf{u} = \frac{-ac\Delta^2 + bdf_1^2 + if_1(ad - bc)\Delta}{a^2\Delta^2 + b^2f_1^2} \quad (29)$$

As the epipoles were chosen on the u axis (cf. equation (22)) and the conic σ_1 is a circle centred in the origin, the polar line will be parallel to the v axis, and \mathbf{p}_1 and \mathbf{p}'_1 will be symmetrical with respect to the u axis. It follows that:

$$(\mathbf{m} + i\mathbf{u}) = -(\mathbf{m} - i\mathbf{u}) \Rightarrow \mathbf{m} = 0 \quad (30)$$

Imposing this condition into equation (29), f_1 can be determined as:

$$f_1 = \sqrt{\frac{-ac}{ac+bd}} \quad (31a)$$

With similar reasoning, the focal length of the second camera, f_2 , can be computed as:

$$f_2 = \sqrt{\frac{-ab}{ab+cd}} \quad (31b)$$

This solution is allowed as far as the two optical axes do not intersect. When they do, the principal points of the two cameras, \mathbf{c}_1 and \mathbf{c}_2 , correspond to each other in the epipolar transformation (equation (10)). In this situation, the following relationship holds:

$$\mathbf{c}_2^T \mathbf{F} \mathbf{c}_1 = 0 \Rightarrow [0, 0, w_2]^T \mathbf{F} [0, 0, w_1] = 0 \quad (32)$$

from which it can be seen that the four elements of \mathbf{F} in equation (23) are linearly dependent. It follows that f_1 and f_2 can be determined only up to a scale factor. This condition suggests a certain care in the setting-up of the cameras although this condition is practically impossible in real applications.

2.4 The essential matrix and the external parameters

Once the matrices \mathbf{K}_1 and \mathbf{K}_2 have been determined, the fundamental matrix \mathbf{F} can be transformed into the essential matrix \mathbf{E} which contains only the external parameters. An elegant algebraic solution to determine \mathbf{R} and \mathbf{T} is based on the Singular Value Decomposition (svd) of the matrix $\mathbf{E}^{(12,24)}$. Being the right kernel of \mathbf{E} equal to $\mathbf{T} \wedge$ (equations (13) and (16)), the product $\mathbf{E}\mathbf{T}$ is equal to zero. It follows that the vector \mathbf{T} is parallel to the eigenvector associated to the smallest eigenvalue, w_3 , of \mathbf{E} . Expressing the essential matrix in term of its svd:

$$\mathbf{E} = [\mathbf{U} \mathbf{W} \mathbf{V}^T] \quad (33)$$

the relative position \mathbf{T} can be determined up to the sign and to a scale factor, as the two versors:

$$\mathbf{T}_1 = \mathbf{V}^3 \quad (34a)$$

$$\mathbf{T}_2 = -\mathbf{V}^3 \quad (34b)$$

Both equations (34) satisfy equation (16). \mathbf{R} can assume one of the two following forms:

$$\mathbf{R}_1 = \mathbf{U} \mathbf{Z} \mathbf{V}^T \quad (35a)$$

$$\mathbf{R}_2 = \mathbf{U} \mathbf{Z}^T \mathbf{V}^T \quad (35b)$$

with:

$$\mathbf{Z} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (36)$$

which both satisfy equation (16): equations (35) allow to determine \mathbf{R} up to rotations of 180 degrees around the rotation axis. From these four possible solutions the correct matrices, \mathbf{R} and \mathbf{T} , are obtained by constraining the reconstructed position of the 3D control points to be in front of the two cameras^(11,24).

It should be remarked that \mathbf{T} is obtained up to a multiplicative factor. This reflects the fact that the 3D world can be reconstructed through the epipolar constraint only up to a scale factor^(2,8,11): a large scene seen from a large distance by a stereo-system with a large inter-camera distance and a smaller scene seen from a smaller distance by a stereo-system with a smaller inter-camera distance, have the same image on the target of the two cameras.

2.5 Determination of the 3D scale factor

In the solution through the essential matrix, $\|\mathbf{T}\| = 1$ has been assumed (equation (34)). The inter-camera distance is therefore taken as the norm of the 3D space. The 3D points position will be scaled accordingly to this choice:

$$k = \|\mathbf{T}^t\| \quad \mathbf{T}^t = k \mathbf{T} \quad \mathbf{P}^t = k \mathbf{P} \quad (37)$$

where \mathbf{P}^t and \mathbf{T}^t are the true 3D position of a point and the true inter-camera distance. \mathbf{P} is the 3D position reconstructed with $\|\mathbf{T}\| = 1$. To determine the value of k , a possibility is to survey a segment whose length is precisely known. Whenever control points are used, a pair of points, whose distance is precisely known, can be used instead of the segment. Let us call these two points \mathbf{P}_1 and \mathbf{P}_2 , it follows:

$$(\mathbf{P}_1^t - \mathbf{P}_2^t) = k (\mathbf{P}_1 - \mathbf{P}_2) \Rightarrow d^t = k d \Rightarrow k = d^t/d \quad (38)$$

However, in the real situation, due to noise on the target coordinates, the 3D reconstruction of \mathbf{P}_1 and \mathbf{P}_2 will generally be not exact. As a consequence, the distance, d , will not be precisely reconstructed and the value of k will be erroneous. To avoid this, multiple measurements are required and the scale factor can be determined as the mean value of a set of M distances between M different pairs of control points:

$$k = E\left(\frac{d^t}{d}\right) = \frac{1}{M} \sum_{j=1}^M \frac{d_j^t}{d_j} \quad (39)$$

In particular, when the control points are the extremes of a single rigid bar, equation (39) becomes:

$$k = \frac{d^t}{M} \sum_{j=1}^M \frac{1}{d_j} \quad (40)$$

3. Determination of the two principal points

Up to now, the position of the two principal points was supposed known. When this is not the case, we show here that it can be determined adjusting the position of the principal points to minimize a figure of merit derived from geometrical considerations on the 3D reconstruction of the control points (indirect estimation).

3.1 Indirect estimation

When the two principal points are displaced with respect to their true value, the offset coordinates (equation (3)) will be also displaced and the estimate of the focal lengths and of the external parameters will be inaccurate. This reflects into an error in the reconstruction of the bar length which will be minimum when the parameters have been correctly estimated. The mean absolute error on the bar length, defined as:

$$e_d = \frac{1}{M} \sum_{j=1}^M |d_j - d^t| \quad (41)$$

will be considered here. Another geometrical quantity affected by the displacement of the principal points is the distance between the two straight lines through the perspective centers and the 2D measured points ($\overline{\mathbf{p}_1\mathbf{C}_1}$ and $\overline{\mathbf{p}_2\mathbf{C}_2}$ in Fig. 1). In the ideal case, these lines do intersect, but, in practice, due to measurement errors on \mathbf{p}_1 and \mathbf{p}_2 , they will be generally skewed. As a consequence, equation (10) will not be exactly satisfied and it should be rewritten as:

$$\mathbf{p}_1^T \mathbf{F} \mathbf{p}_2 = e_e \quad (42)$$

where e_e is the coplanarity error. This error has been widely used to solve the correspondence problem^(25,26) and it has been recently proposed in the calibration framework⁽²⁷⁾.

These considerations have led us to formulate the estimate of the internal parameters as a non-linear minimization problem of the following cost function:

$$\min_{(\mathbf{c}_1, \mathbf{c}_2)} \left(\mathbf{a} \frac{1}{N/2} \sum_{j=1}^{N/2} |d_j - d^t| + \mathbf{b} \frac{1}{N} \sum_{i=1}^N |\mathbf{p}_{i1} \mathbf{F} \mathbf{p}_{i2}| \right) \quad (43)$$

where N is the number of control points. The second term does not allow to determine the right norm of the 3D space (equation (10)) and it can be used only as a refinement of the solution which, in its bulk part, is obtained through the first term. Therefore $\mathbf{b} \ll \mathbf{a}$ should hold. In the following $\mathbf{a} = 1$ and $\mathbf{b} = 0.1$ have been adopted.

3.2 Evolutionary optimisation and the principal points

The above formulation leads to a highly non-linear implicit function of \mathbf{c}_1 and \mathbf{c}_2 . These in fact influence the value of the other calibration parameters which, in turn, influence the value of the 3D points position and therefore of the cost function. The global minimum of equation (42) is therefore hard to be found with classical gradient-based methods. In this paper a solution offered by evolutionary optimisation, which has proved particularly successful in the solution of difficult optimisation problems, is exploited⁽²⁸⁾. In this domain, the search for the optimal solution is guided by the evaluation of the cost function alone. The single solution is represented as an individual, the ensemble of possible solutions as a population, and the cost associated to each solution as the degree of fitness of the population element associated to the solution.

Before starting the search for the optimal pair of principal points, a search region is defined. Taking into account that the maximum offset of the principal points with respect to the image center is $\pm T_{\max}$, a hyperbox search region, H , is created in the four-dimensional space of the two principal points ($[\mathbf{c}_1 \mid \mathbf{c}_2]$):

$$H = \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \end{bmatrix} \pm T_{\max} \quad (44)$$

The optimisation strategy is initialised generating inside H , a random set of NP pairs of principal points, $FA = \{[\mathbf{c}_{1i} \mid \mathbf{c}_{2i}] \in H\}$, called “fathers” of the “population”. For these points, the external parameters, the focal lengths and the 3D scale factor is determined following the procedures reported in Sections 2.3, 2.4 and 2.5; and the “fitness” is determined through equation (43) for each father. Notice that for some of the fathers, an imaginary value of the focal lengths may come out from equations (31). This happens when the principal points are far-off the true position. In this case a high cost is associated to the related fathers. After the fitness has been computed for all the fathers a second set, S , of NP pairs of principal points, called “sons”, is obtained by “mutating” the set FA . The fitness of the sons is evaluated through equation (43) as for the fathers. In the mutation process, each father is displaced by a random quantity, $\mathbf{z}_i(0, \sigma)$,

where \mathbf{z}_i is a Gaussian variable in the four dimensional space with zero mean and standard deviation σ , obtaining:

$$[\mathbf{c}_{1i} | \mathbf{c}_{2i}]^{k+1} = [\mathbf{c}_{1i} | \mathbf{c}_{2i}]^k + h^k \mathbf{z}_i(0, \sigma) \quad (45)$$

where k indicates the k^{th} generation. σ and h^k concur in the determination of the amplitude of the displacement for each father.

The fitness of each son is compared with that of the father from which it was generated and the one of the two elements which exhibits the best fitness will serve as father for the next “generation”. In more evolutionistic terms: the father and his son are placed in competition for survival and the selection eliminates the weaker for the next generation⁽²⁹⁾, that is the one which gives the poorer solution. The process of generating new sons and selecting those which have the best fitness is iterated until the fitness does not increase anymore.

To increase the resolution in the determination of the principal points, a deterministic annealing scheduling⁽³⁰⁾ of the search region amplitude has been introduced by setting h^k as:

$$h^k = \frac{h^{\text{initial}}}{\sqrt{\ln(k+1)}} \quad (46)$$

With this setting, the amplitude of the search region (equation (44)) decreases with the generations, allowing a thicker sampling of the best region. The role of σ is to restrict the search region if a meaningful number of sons have given a better fitness, or to enlarge it, if the fathers were fitter. For this purpose, the value of σ is set according to the ratio, r , between the number of winning sons and winning⁽²⁹⁾ as follows:

$$\begin{aligned} r < 1/5 & \quad \mathbf{s}^{k+1} = \mathbf{s}^k / c \\ r = 1/5 & \quad \mathbf{s}^{k+1} = \mathbf{s}^k \\ r > 1/5 & \quad \mathbf{s}^{k+1} = \mathbf{s}^k c \end{aligned} \quad (47)$$

with $c = 0.85^{1/D}$ where D is the dimension of the solution space, in our case $D=4$.

To speed up the convergence and to avoid bias in the results due to the particular set of calibration points adopted, not all the control points are used at the same time. At each generation, a different sub-set of points is extracted and used for calibration and a different sub-set is extracted and used for evaluating the fitness. This procedure is commonly used in statistics and it goes under the name of bootstrap⁽³¹⁾.

4. Summary of the calibration procedure

The entire calibration procedure can be summarized as follows (Fig. 3):

1. A set of NC points, $\{\mathbf{P}^{\text{NC}}_c\}$, extremes of a rigid bar is surveyed by a pair of cameras and their position on the image planes are measured as $\{\mathbf{p}^{\text{NC}}_{1c}\}$ and $\{\mathbf{p}^{\text{NC}}_{2c}\}$. These constitute the set of the control points.
2. A population of NP pairs of principal points $\text{FA} = \{\mathbf{c}_1 \mid \mathbf{c}_2\}$ is randomly generated inside the search area H. These constitute the fathers of the population.
3. The reference fitness is set to 0.
4. A subset of M ($< \text{NC}/2$) pairs of control points: $\{\mathbf{P}_c\}$, $\{\mathbf{p}_{1c}\}$ and $\{\mathbf{p}_{2c}\}$, is extracted from the set $\{\mathbf{P}^{\text{NC}}_c\}$. This will be used to estimate the focal lengths, the external parameters and the object scale factor.
5. A second subset of T ($< \text{NC}/2$) pairs of points: $\{\mathbf{P}^{\text{T}}_c\}$, $\{\mathbf{p}_{1\text{T}}\}$ and $\{\mathbf{p}_{2\text{T}}\}$ is extracted from the set $\{\mathbf{P}^{\text{NC}}_c\}$. This will be used to determine the fitness and it will serve as test set.
6. A set of NP fundamental matrices, $\{\mathbf{F}\}$, are computed from $\{\mathbf{p}_{1c}\}$ and $\{\mathbf{p}_{2c}\}$, one for each element of FA.
7. A set of NP pairs of focal lengths, $\{f_1\}$ and $\{f_2\}$, are determined as reported in Section 2.3. When an imaginary value results for the focal length, a low value of fitness is associated to the related pair of fathers.
8. The coordinates of $\{\mathbf{p}_{1c}\}$ and $\{\mathbf{p}_{2c}\}$ are normalised through $\{f_1\}$, $\{f_2\}$ (equation (5)). A set of NP essential matrices, $\{\mathbf{E}\}$, is computed.
9. A set of NP relative orientations, $\{\mathbf{R}\}$, and normalised relative positions, $\{\mathbf{T}\}$ (with $\|\mathbf{T}\| = 1$), are computed from the matrices $\{\mathbf{E}\}$ according to Section 2.4.
10. A sub-set of NC 3D scale factors, $\{k\}$, is determined.
11. The 3D position of the test points is reconstructed through the estimated parameters $\{\mathbf{R}\}$, $\{\mathbf{T}\}$, $\{k\}$, $\{f_1\}$, $\{f_2\}$ and FA, originating NP sets of 3D points constituted of NT points each.
12. The cost function in equation (43) (fitness) is computed for each set of 3D test points.
13. A new population of NP pairs of principal points, the sons, SA, is generated inside the search area H by random mutation of the fathers set.
14. The internal and external parameters, and the fitness value associated with the sons is determined according to the steps 6-12.

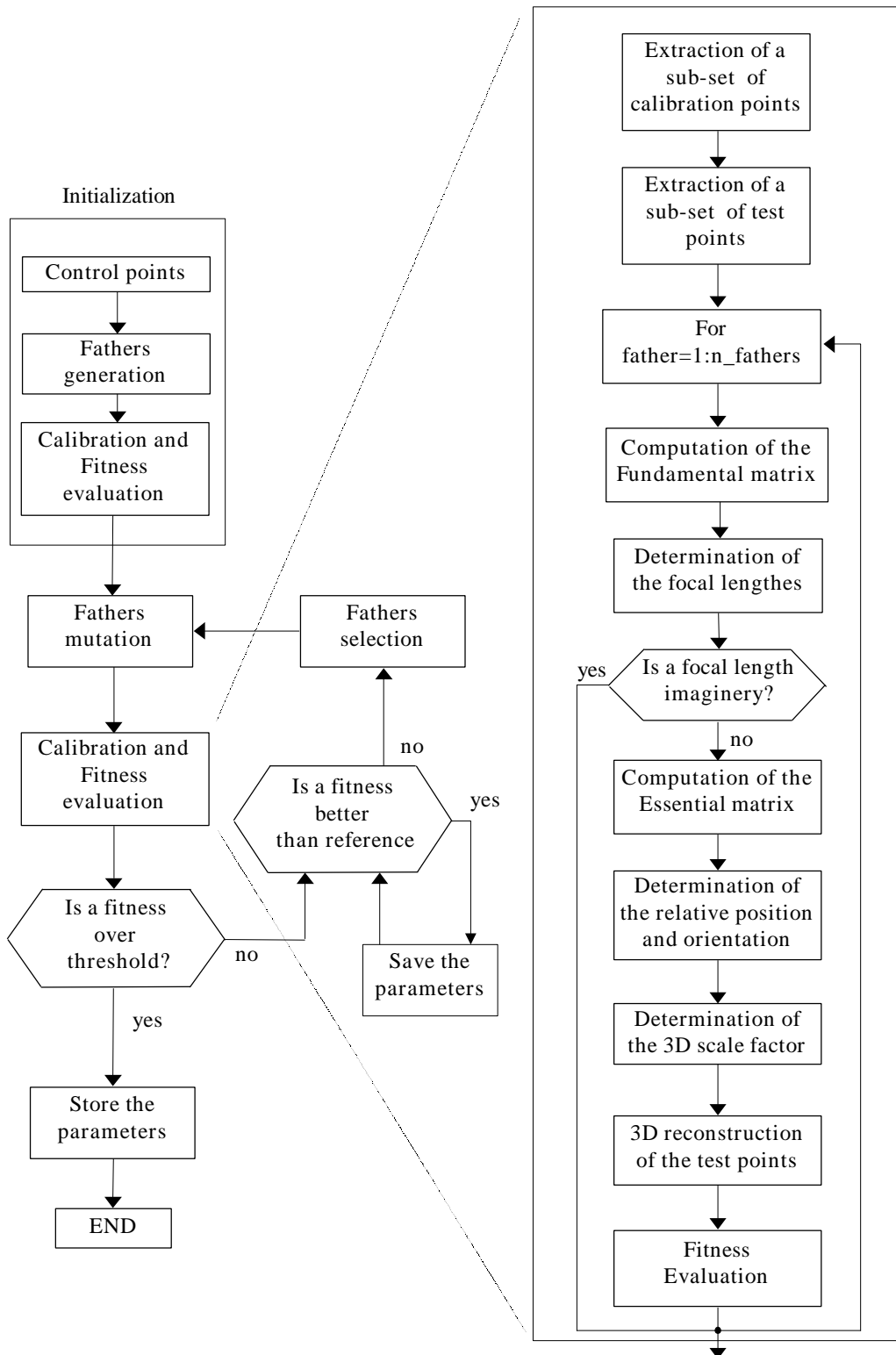


Figure 3. The flow chart of the calibration procedure is reported on the left. On the right the flow chart of the block calibration and fitness evaluation is exploded in its main components.

15. If the cost function is lower than the actual reference fitness, the reference fitness is updated and the corresponding parameters are saved (cf. Fig. 3, block diagram on the left).

16. The one between each father and its son which has the best fitness wins and it will serve as father in the next generation. With these new fathers, steps 4-12 are repeated.

The search for the principal points (and for the other parameters) ends when the cost function drops under a predefined threshold or it does not decrease anymore.

5. Experimental results

The accuracy obtained with this calibration method has been evaluated on real data and compared with that obtained through a calibration methods based on control points of known 3D coordinates, based on the iterative solution of the perspective equations (ILSSC)^(4,10,32).

To calibrate, a rigid bar is moved inside the working volume while it is surveyed by a pair of video cameras. A small passive marker is placed on each of its extremities. The markers coordinates on the image plane of the two cameras are measured automatically by the Elite system⁽³³⁾. This recognises the markers thanks to a template matching algorithm implemented on a custom VLSI chip⁽³⁴⁾. The correspondence between the points on the two cameras has been carried out automatically through the Smart3D tracking system⁽³⁵⁾. With this procedure, a very large amount of calibration points, spread inside the working volume, can be collected in a very few time.

To increase the accuracy, the main distortions component, which is due to the different scales on the two target axes⁽²⁾, has been corrected multiplying the horizontal target coordinate by a shrinkage factor obtained through a 2D in-house calibration. This factor, which for the Elite system is 1.4779, is a characteristic of the cameras as it depends only on the electronics and on the target dimensions, and it does not depend on the lenses. Therefore the same shrinkage factor applies to different optics, different apertures and different focusing of the cameras. The same in-house calibration allows to determine the search region H (equation (44)), as a hyperbox centered in the point [(128, 128) | (128,128)] with side $T=\pm 45$ pixels.

Two experiments have been carried out whose quantitative results are reported in Tables I and II. In the first experiment, a zoom lens has been mounted on the cameras (focal length approximately 30mm) and a bar carrying two markers 99.1mm apart was used. In the second one a pair of wide-angle lenses with focal length 8.5mm and a bar carrying two markers 199.8mm apart were adopted. The two cameras were positioned approximately at the same height with a relative orientation of 60 degrees.

Two sets of control points were acquired for each experiment. The first set, S_1 , was obtained surveying the bar in motion inside the calibration volume for 40 seconds collecting a total of 4,000 matched pairs of 2D points for each camera. After discarding those frames where the two markers were not visible on both cameras, a total of $NC=2749$ and $NC=2530$ calibration points, respectively, were left for calibration. From this set of data points, a sub-set of $M=150$ pairs of calibration points and a sub-set of $NT=100$ pairs of test points were randomly extracted at each optimisation step of the calibration procedure. The second set of control points, S_2 , is constituted of the markers positioned on a planar grid surveyed in three known parallel positions inside the working volume. For the first experiment, each grid carried 30 markers (5x6) with an inter-marker distance of 50mm and an inter-plane distance of 200mm; in the second experiment, each grid carried 56 markers (7x8) with an inter-marker distance of 150mm and an inter-planar distance of 400mm. The cameras set-up was the same used for the acquisition of the bar.

The accuracy was quantitatively assessed as follows:

- a) The 3D coordinates of the markers on the bar extremities (set S_1 of control points) are reconstructed with the parameters determined through ILSSC and through bar calibration. From the 3D coordinates of the markers, their 3D distance is determined frame by frame obtaining the set of distances D_1 and D_2 respectively. The difference between the true bar length and D_1 or D_2 is computed.
- b) The 3D coordinates of the points on the grids (set S_2 of control points) are reconstructed with the parameters determined through ILSSC and through bar calibration. These coordinates cannot be directly compared because the reference systems are different: for ILSSC it is solid with the calibration grids while for bar calibration it is solid with one of the two cameras. Relative measurements have therefore been adopted: the error on the X and Y directions is computed as the mean and standard deviation of the error on the distance between each pair of two consecutive markers on the horizontal and vertical marker lines on the calibration grid. The error on the Z direction is computed as the mean and standard deviation of the error on the distance between each pair of markers in the same position on two consecutive parallel grids.

As can be seen in Table I and II, and in Fig. 4, bar calibration gives an accuracy which is equal or even slightly better with respect to ILSSC both on the bar length and on the relative distances on the grids. An example of the sequence of the principle points examined by the evolutionary optimization is reported in Fig. 5a and 5b. Although the search is carried out in the whole search region, H (equation (44)), the ensemble of the principle points is driven into a small

target region which includes the true position of the two principal points. This assures a high resolution in the estimate. Moreover, as the region H is explored in parallel, using MP elements at each step, the error decreases very fast. The resulting calibration parameters are adequate already at the first step when high accuracy is not required. As can be seen in Fig. 5c, for the best pair of principal points the mean error is only 6.5 times the final error on the grid and 1.7 times on the bar length. The total time required by the forty optimisation steps of the algorithm is 74s on a Pentium MMX, 200Mhz.

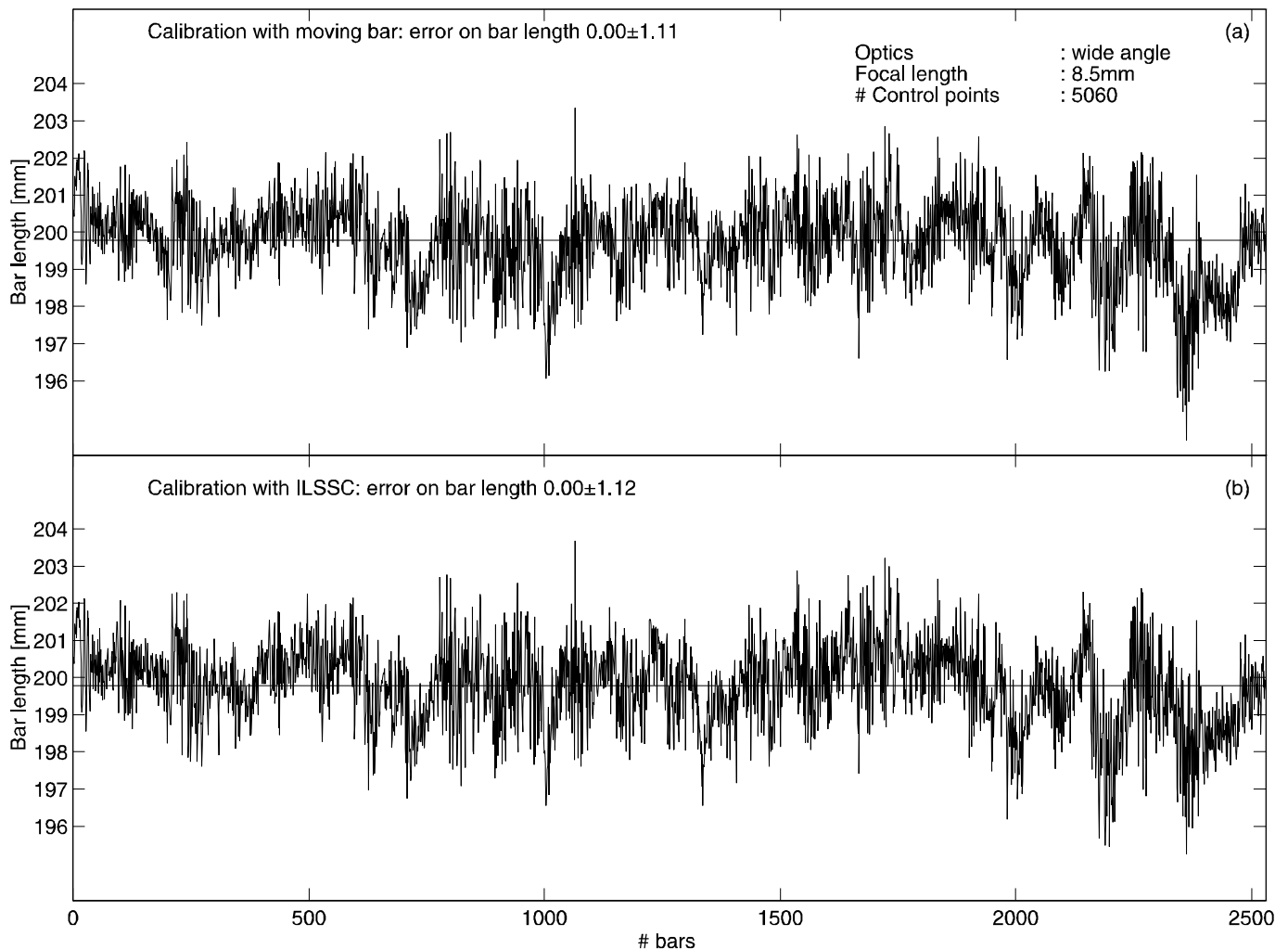


Figure 4. The reconstructed distances between the two markers on the bar extremes frame by frame. They have been computed from the parameters obtained with bar calibration (Figure 4a) and with ILSSC (Figure 4b). The data are referred to the second experiment (wide-angle lens with focal length = 8.5mm).

6. Discussion and Conclusion

The only metric information required by bar calibration is the true 3D distance between the bar extremes, and this is in fact a critical parameter. When the distance is not given correctly, a

degradation in the accuracy is introduced. However, as this measurement is taken only once, it is reasonable to assume that it can be performed with an adequate accuracy.

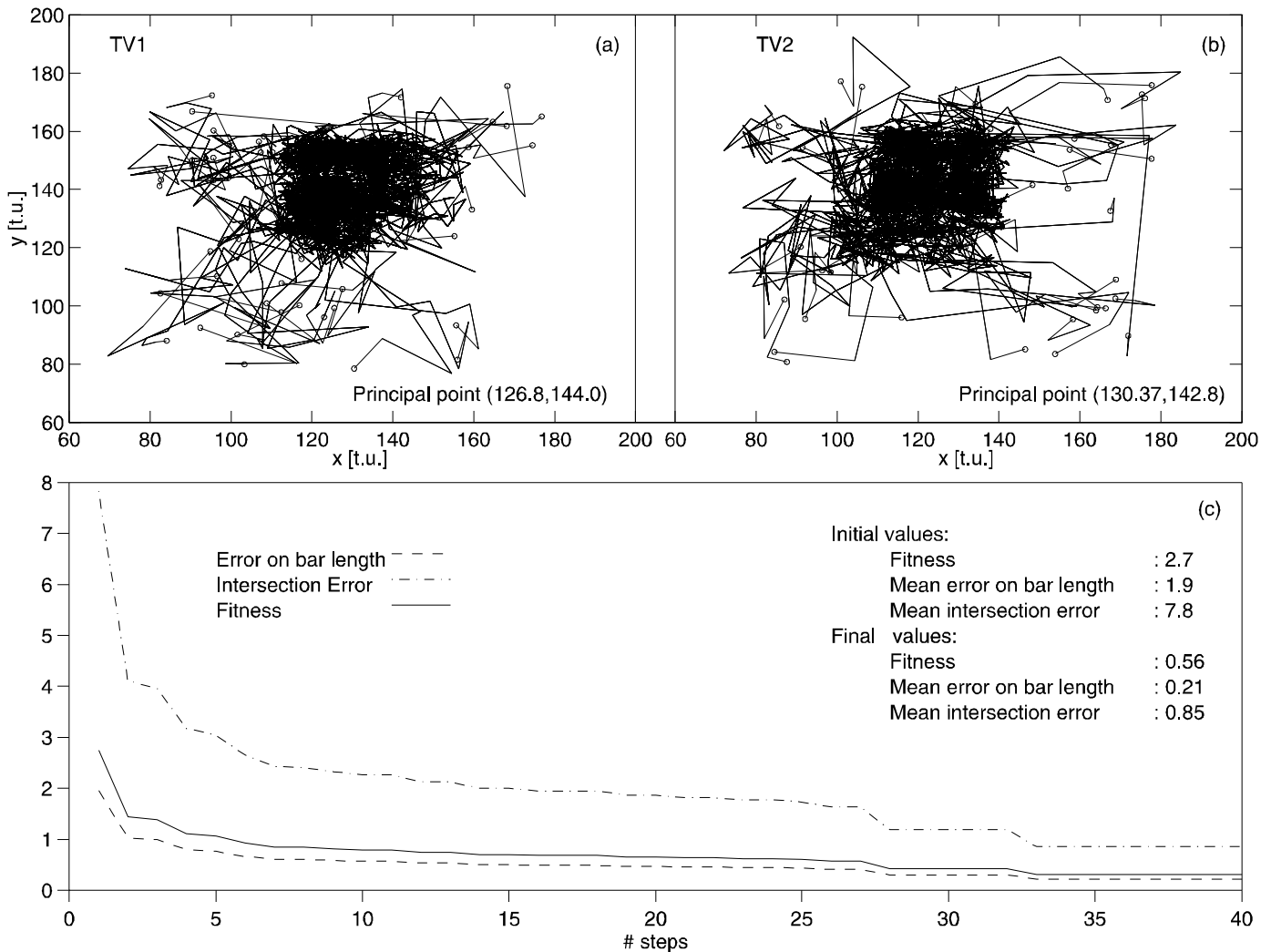


Figure 5. The trajectory of the principal points is reported in Figure 5a for TV1 and in Figure 5b for TV2. The initial position of each pair of principal points is plotted as empty circles and the final one by filled circles. The 2D coordinates are expressed in target units (t.u.). In Figure 5c, the fitness, the mean error on the bar length and the mean intersection error are reported for each step of the calibration in experiment 1.

Another critical parameter is the angle between the optical axes of the two cameras. When it departs from 90 degrees, the accuracy decreases with power law up to 6 times for angles of 20 or 160 degrees (e.g. Fig. 3 of Ref. (32))^(16,32,36). This can be a problem for bar calibration where the 3D reconstruction is an essential part of the calibration procedure while it was not for ILSSC where each camera is calibrated separately. Therefore when the relative angles are very small or very large, many control points are required to achieve a good estimate.

When the stereo-pair has been properly set-up and the length of the bar is measured with adequate precision, the accuracy of bar calibration is comparable and even superior to that

obtained with points of known 3D coordinates. This allows to substitute a simple rigid bar to the cumbersome calibration structures required by classical calibration methods and it suggests to introduce bar calibration as a standard tool for those stereo systems which have to be calibrated before use.

Table I

The accuracy obtained with ILSSC and bar calibration when the cameras were equipped with zoom lens (focal length $\cong 30\text{mm}$). The data from two different calibrations by using the bar are indicated as Bar, 1st and 2nd calibration.

Bar length: 99.1 mm. Grid contains 30 markers (5×6), surveyed in three parallel positions 200mm apart. Distance between two consecutive markers on the grid: 50mm. Calibrated volume $0.25 \times 0.20 \times 0.40 \text{ m}^3$. Working volume $0.6 \times 0.9 \times 0.9 \text{ m}^3$. Evolutionary optimization: 100 elements and 40 generations.			
	ILSSC	Bar, 1 st calibration	Bar, 2 nd calibration
bar, mean error	0.00±0.50	0.00±0.25	0.00±0.21
grid, x error	0.04±0.84	0.00±0.81	0.02±0.81
grid, y error	0.02±0.07	0.05±0.02	0.04±0.02
grid, z error	0.07±0.61	0.15±0.58	0.00±0.58

Table II

The accuracy obtained with ILSSC and bar calibration when the cameras were equipped with wide-angle lens (focal length = 8.5mm). The data from two different calibrations by using the bar are indicated as Bar, 1st and 2nd calibration.

Bar length: 199.8 mm. Grid contains 56 markers (7×8), surveyed in three parallel positions 400mm apart. Distance between two consecutive markers on the grid: 150mm. Calibrated volume $0.9 \times 0.9 \times 0.8 \text{ m}^3$. Working volume $1.3 \times 1.2 \times 1.5 \text{ m}^3$. Evolutionary optimization: 100 elements and 40 generations.			
	ILSSC	Bar, 1 st calibration	Bar, 2 nd calibration
bar, mean error	0.00±1.12	0.00±1.12	0.01±1.16
grid, x error	0.26±0.93	0.47±1.08	0.56±0.93
grid, y error	0.18±0.07	0.15±0.02	0.08±0.04
grid, z error	0.59±2.94	0.60±2.99	0.52±2.99

Appendix A: What is represented in the fundamental matrix?

We summarise here the considerations of projective geometry which limit to seven the number of independent parameters which can be determined using the epipolar constraint. From equation (10) it can be seen that Fundamental matrix \mathbf{F} is not of full rank. Let $\mathbf{e}_1 = [u_{e_1}, v_{e_1}, w_{e_1}]$ and $\mathbf{e}_2 = [u_{e_2}, v_{e_2}, w_{e_2}]$ be the two epipoles. Let us now consider a point \mathbf{p}_1 onto the image plane of the first camera. The epipolar line \mathbf{l}_{p1} to which \mathbf{p}_1 belongs (cf. Fig. 1) has homogeneous representation $[l_{u1}, l_{v1}, l_{w1}]^T = (\mathbf{e}_1 \wedge \mathbf{p}_1)^{(3)}$. The transformations $\mathbf{p}_1 \rightarrow \mathbf{l}_{p1}$ and $\mathbf{p}_2 \rightarrow \mathbf{l}_{p2}$ are therefore linear and they can be expressed as:

$$\mathbf{l}_{pj} = [l_{uj}, l_{vj}, l_{wj}]^T = \mathbf{Q}_j \mathbf{p}_j \quad (\text{A1})$$

where \mathbf{Q}_j is a 3x3 matrix of rank two, function of the epipole coordinates $[u_{ej}, v_{ej}, w_{ej}]$, two of which are independent:

$$\mathbf{Q}_j = \mathbf{e}_j \wedge = \begin{bmatrix} 0 & -w_{e_j} & v_{e_j} \\ w_{e_j} & 0 & -u_{e_j} \\ -v_{e_j} & u_{e_j} & 0 \end{bmatrix} \quad (\text{A2})$$

We now search for the transformation between two corresponding epipolar lines on the two image planes. This is defined by a 3x3 matrix, \mathbf{G} , defined up to scale factor, such that:

$$\mathbf{l}_{p2} = \mathbf{G} \mathbf{l}_{p1} \quad (\text{A3})$$

Noticing that $\mathbf{e}_2^T \mathbf{l}_{p2} = 0$, it follows:

$$\mathbf{e}_2^T \mathbf{G} \mathbf{Q}_1 \mathbf{p}_1 = 0 \quad (\text{A4})$$

Being $\mathbf{Q}_1 \mathbf{p}_1 \neq 0$ (equation (A1)), it follows that the right kernel of \mathbf{G} must be equal to \mathbf{Q}_2^T up to a scale factor⁽³⁷⁾:

$$\mathbf{G} = \mathbf{Q}_2^T \mathbf{J} \quad (\text{A5})$$

and the following relationships between \mathbf{p}_1 and \mathbf{p}_2 and \mathbf{l}_{p1} and \mathbf{l}_{p2} hold:

$$\mathbf{p}_2^T \mathbf{Q}_2^T \mathbf{J} \mathbf{Q}_1 \mathbf{p}_1 = 0 \quad \mathbf{F} = \mathbf{Q}_2^T \mathbf{J} \mathbf{Q}_1 \quad (\text{A6a})$$

$$\mathbf{l}_{p2}^T \mathbf{J} \mathbf{l}_{p1} = 0 \quad (\text{A6b})$$

\mathbf{J} represents a projective mapping between two corresponding epipolar lines⁽³⁾. As this transformation has the property to preserve the cross-ratio, it will be uniquely identified by four parameters (a,b,c,d) , three of which are independent. Every pair of epipolar line can be identified as:

$$t_1 = -\frac{at_2 + c}{bt_2 + d} \quad (\text{A7})$$

where:

$t_1 = \frac{l_{u1}}{l_{v1}}$ and $t_2 = \frac{l_{u2}}{l_{v2}}$ are the projective coordinates of the two corresponding epipolar lines (l_{w1} and l_{w2} have been assumed equal to one). It follows that the matrix \mathbf{J} assumes the following expression:

$$\mathbf{J} = \begin{bmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (\text{A8})$$

and, from equation (A6a), the Fundamental matrix can be written as:

$$\mathbf{F} = \begin{bmatrix} d & -c & -(du_{e1} - cv_{e1}) \\ -b & a & -(av_{e1} - bu_{e1}) \\ -(du_{e2} - bv_{e2}) & -(av_{e2} - cu_{e2}) & v_{e1}(av_{e2} - cu_{e2}) + u_{e1}(du_{e2} - bv_{e2}) \end{bmatrix} \quad (\text{A9})$$

Thus the epipolar transformation can be defined by the four independent coordinates of the two epipoles and the three parameters which define the bilinear transformation between corresponding epipolar lines for a total of seven independent parameters.

As the right hand side of \mathbf{F} is the matrix $\mathbf{Q}_1 = (\mathbf{e}_1 \wedge)$, the product $\mathbf{F}\mathbf{e}_1$ is equal to zero (the same is true for \mathbf{e}_2). This represents the fact that the epipoles have not a corresponding epipolar line on the other camera.

References

1. P. R. Wolf, *Elements of photogrammetry*. McGraw-Hill, New York (1983).
2. J. Weng, P. Cohen P. and M. Henriou Camera calibration with distortion models and accuracy evaluation. *IEEE Trans. Patt. Anal. Machine Intell.*, **14**(10), 965-979 (1992).
3. O. D. Faugeras, *Three-Dimensional computer vision*. MIT Press, Cambridge, MA (1993).
4. J. Weng, T. S. Huang and N. Ahuja, Motion and structure from two perspective views: algorithms, error analysis and error estimation. *IEEE Trans. Patt. Anal. Machine Intell.*, **11**(11) 451-476 (1989).
5. J. F. Kenefick, Ultra-precise analytics. *Photogramm. Engng.*, **37**, 1167-1187 (1971).
6. A. Gruen and H. A. Beyer, System calibration through self-calibration, Invited paper, *Workshop on Camera calibration and orientation in computer vision, XVIIth ISPRS Congress*, Washington D.C. (1992).
7. S. Ullman, *The Interpretation of Visual Motion*. Cambridge, MA, M.I.T. Press, (1979).
8. J. Dapena, A. H. Everett and J. A. Miller Three-dimensional cinematography with control object of unknown shape. *J. Biomech.* **15**, 11-19 (1982).
9. B. K. Horn, *Robot Vision*. MIT Press (1986).
10. J. Weng, T. S. Huang and N. Ahuja Optimal motion and structure estimation. *IEEE Trans. Patt. Anal. Machine Intell.* **15**(9), 864-884 (1993).
11. H. C. Longuet-Higgins, A computer algorithm for reconstructing a scene from two projections. *Nature*, **293**, 133-135 (1981).
12. R. I. Hartley, Estimation of relative positions for uncalibrated cameras. *Computer vision - ECCV '92, LNCS-Series*, Springer-Verlag, Berlin, **588**, 579-587 (1992).
13. R. I. Hartley, Projective reconstruction and invariants from multiple images. *IEEE Trans. Patt. Anal. Machine Intell.*, **16**(10) 1036-1041 (1994).
14. O. D. Faugeras, Q. T. Luong and S. J. Maybank, Camera self-calibration: Theory and experiments. In *Computer vision - ECCV '92, LNCS-Series*, Springer-Verlag, Berlin, **588**, 321-334 (1992).
15. S. J. Maybank and O. D. Faugeras, A theory of self-calibration of a moving camera. *Int. Journal of Comp. Vis.*, **8**(2), 123-151 (1992).
16. Z. Zhang Determining the Epipolar Geometry and its Uncertainty: A Review, *Int. J. Comp. Vision*, **27**(2), 161-195 (1998).
17. N. A. Borghese and P. Perona, Calibration of a stereo system with points of unknown location. *Proc. XIVth Int. Conf. Soc. Biomech. ISB*, Paris, 202-203 (1993).
18. R. K. Lenz and R. Y. Tsai, Techniques for calibration of the scale factor and image center for high accuracy 3D machine vision metrology, *IEEE Trans. Patt. Anal. Machine Intell.* **10**(5) 713-720 (1988).
19. G. Wei G. and S. D. Ma, Implicit and explicit camera calibration theory and experiments. *IEEE Trans. Patt. Anal. Machine Intell.*, **16**(5), 469-480 (1994).

20. R. Tsai, A versatile camera calibration technique for high accuracy 3D machine vision metrology using off-the-shelf TVcameras and lenses. *IEEE J. Robotics Automat.* **3**(4), 35-47 (1987).
21. S. Shah and J. K. Aggarwal, Intrinsic parameter calibration procedure for a (high distortion) fish-eye lens camera with distortion model and accuracy estimation. *Pattern Recognition* **29**(11), 1775-1788 (1996).
22. Q. T. Luong and O. D. Faugeras, The fundamental matrix: theory, algorithms, and stability analysis. *Int. Journal of Computer Vision*, **17**(1), 43-76 (1996).
23. O. D. Faugeras, Stratification of three-dimensional vision: projective, affine, and metric representations. *J. Opt. Soc. Am. A*, **12**(3) 465-484 (1995).
24. S. Soatto, P. Perona, R. Frezza and G. Picci, Recursive motion and structure estimation with complete error characterization, *ECCV '92, LNCS-Series*, Springer-Verlag,, Berlin, **588**, 428-433 (1993).
25. N. A. Borghese, SMART3D: Tracking of movements surveyed by a multiple set of TVcameras. *Proc. X Symp. ISBS*, 108-111 (1992).
26. X. Hu and N. Ahuja, Matching Point Features with Ordered Geometric, Rigidity, and Disparity Constraints, *IEEE Trans. Patt. Anal. and Machine Intell.* **16**(10), 1041-1049 (1994).
27. R. I. Hartley, In defense of 8-point algorithm, *IEEE Trans. Patt. Anal. and Machine Intell.* **19**(6), 580-593 (1997).
28. D. B. Fogel, An introduction to simulated evolutionary optimization, *IEEE Trans. Neural Net.*, **5**(1), 3-14 (1994).
29. T. Bäck T., G. Rudolph and H. P. Schwefel, Evolutionary programming and evolution strategies: similarities and differences, *Proc. of the Second Ann. Conf. on Evolutionary Programming*, 11-22 (1993).
30. K. Rose, F. Gurewitz, and G. Fox Statistical mechanics and phase transitions in clustering. *Physical Rev. Letters*, **65**(8), 945-948 (1990).
31. P. J. Huber, *Robust Statistics*, Wiley, NY (1981).
32. N.A. Borghese and G. Ferrigno, An algorithm for 3D automatic movement detection by means of standard TVcameras. *IEEE Trans. Biomed. Eng.*, **37**(12) 1221-1225 (1990).
33. G. Ferrigno and A. Pedotti, ELITE: A digital dedicated hardware system for movement analysis via real-time TV signal processing, *IEEE Trans. Biomed. Eng.*, **32**, 943-950 (1985).
34. N.A. Borghese, M. Di Rienzo, G. Ferrigno and A. Pedotti, Elite: a goal-oriented vision system for moving objects detection, *Robotica*, **9**, 275-282 (1990).
35. N.A. Borghese, SMART3D: Tracking of movements surveyed by a multiple set of TVcameras, *Proc. X ISBS Symposium*, Hermes Edition, Milano, 108-111, (1992).
36. J. Oliensis, Rigorous Bounds for Two-Frame Structure from Motion. *Proc. ECCV '96, LNCS-Series*, Springer-Verlag, Berlin, **1065**, 184-195 (1996).
37. G.H. Golub, C.F. Van Loan, *Matrix Computation*, The John Hopkins University Press, (1991).