

## Linear-Camera Calibration

Radu Horaud, Roger Mohr

LIFIA-IRIMAG  
46, avenue Félix Viallet  
38031 Grenoble France

Boguslaw Lorecki

CYBERNETIX  
rue Albert Einstein - BP 94  
13382 Marseille France

### Abstract

*A method for calibrating linear cameras is described in this paper. We show that the more classical 2-D camera calibration techniques are necessary but not sufficient for solving for the linear-camera calibration problem. We propose a model for linear cameras and we provide a two-step procedure for estimating its parameters. We also show how the extrinsic camera parameters can be made explicit. The accuracy of the calibration method is analysed through an example of application.*

### 1 Introduction

In many applications of computer vision 1-D (linear) cameras may replace 2-D (matrix) cameras. Inspection of parts, for example, is a task that requires, among others, two main constraints: efficiency and accuracy. Linear cameras fit both these constraints for the following reasons. First the 1-D signal they provide is easier and faster to process than 2-D images. Second the current size of a linear camera may reach 4096 pixels. This size is to be compared with the size of a matrix camera which is (at an affordable price) of 512 by 512 pixels. We conclude that linear cameras are well suited for such tasks as inspection that do not involve high-level processing and for which accuracy is a crucial issue.

More generally, the calibration of a camera allows one to establish a relationship between image measurements and 3-D measurements. This relationship is useful for a variety of tasks:

- 3-D reconstruction using either passive stereo or active ranging;
- mobile-robot navigation;
- 3-D object recognition by matching a CAD model to 2-D data;

- hand/eye calibration which allows one to relate measurements in the camera-centered frame to the robot-centered frame, and
- 3-D inspection of parts.

The calibration of a matrix camera has received a lot of attention in the past and robust techniques are now available [6], [5], [2]. Surprisingly enough, at our knowledge linear-camera calibration is not described in the technical literature.

In this paper we show that the currently available camera calibration methods are necessary but not sufficient for solving for the linear-camera calibration problem. Next we propose a two-step camera calibration method that uses the classical approach in conjunction with the *cross-ratio*, a wellknown projective invariant [3]. Finally we define a camera-centered coordinate frame and determine the extrinsic parameters of the camera, that is, the rigid transformation between the calibrating coordinate frame and the camera frame.

### 2 2-D camera calibration

Central projection or the pin-hole model is a good approximation of the geometry of an imaging system. Using central projection the 2-D camera calibration problem is described by the following set of equations [2], [1]:

$$u = \frac{m_{11}X + m_{12}Y + m_{13}Z + m_{14}}{m_{31}X + m_{32}Y + m_{33}Z + m_{34}} \quad (1)$$

$$v = \frac{m_{21}X + m_{22}Y + m_{23}Z + m_{24}}{m_{31}X + m_{32}Y + m_{33}Z + m_{34}} \quad (2)$$

These equations describe the projection of a point defined in a calibrating (world) frame onto the image. The world coordinates of the 3-D point are  $X$ ,  $Y$ , and  $Z$  and the image coordinates of its projection are  $u$

and  $v$ . Notice that these equations have another possible interpretation. They describe the line passing through the center of projection of the camera and through an image point, in world coordinates. The 2-D camera calibration problem is the problem of determining the parameters  $m_{11}$  through  $m_{34}$ . If at least 6 correspondences are available, the calibration problem becomes a linear optimization problem [2]. By a correspondence  $i$  is meant that the set  $\{X_i, Y_i, Z_i, u_i, v_i\}$  is known: the coordinates of a 3-D point are known in world coordinates and its 2-D projection is known as well in image coordinates.

### 3 1-D camera calibration

The best way to describe a linear camera is to think of a matrix camera for which only one line of pixels is active. A 3-D point projects onto this linear image at pixel  $u$  obeying equation (1).

Moreover, in order to be seen by the camera, this point is constrained to lie in the *viewing plane*, the plane defined by the center of projection and by the active line of pixels, e.g., Figure 1. Without loss of generality, the equation of the viewing plane can be written as:

$$X = pY + qZ + r \quad (3)$$

By substituting equation (3) into equation (1), renaming variables, and noticing that equation (1) describing a projection is defined up to multiplicative term, we obtain:

$$u = \frac{n_1 Y + n_2 Z + n_3}{n_4 Y + n_5 Z + 1} \quad (4)$$

Equations (3) and (4) define the linear-camera model, that is, they describe the line passing through the center of projection and an image pixel, in world coordinates. This line is also called the *viewing line*.

This model has 8 parameters associated with it:  $n_1, n_2, n_3, n_4, n_5$ , and  $p, q, r$ . The problem of determining the camera model is the problem of estimating these parameters. In the Appendix we prove that 8 parameters are sufficient for describing a linear-camera model. Hence the linear-camera calibration problem is twofold:

- Estimate the parameters  $n_1, n_2, n_3, n_4, n_5$ . If at least 5 correspondences are available (5 sets  $\{Y_i, Z_i, u_i\}$ ), this is a linear optimization problem similar to the case of 2-D camera calibration.
- Estimate  $p, q, r$ . These parameters can be estimated if and only if world points belonging to the viewing plane are provided.

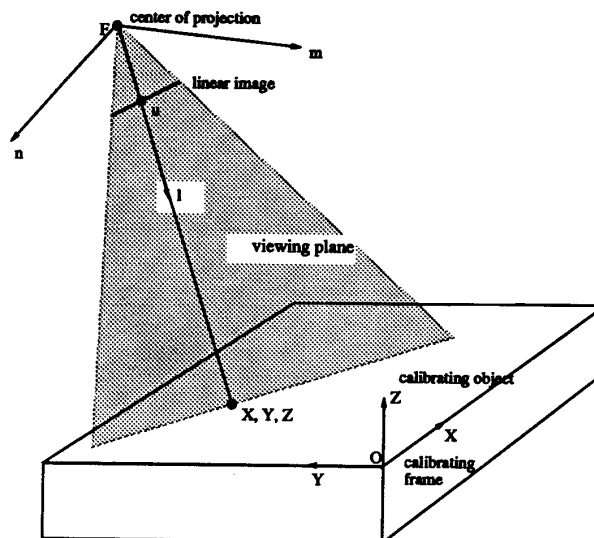


Figure 1: A general view of a linear camera, a calibrating object, and their associated frames

Interesting enough, the linear-camera calibration problem is similar to the problem of calibrating an active range sensor [1]. In the latter case two calibration steps are also necessary: one for calibrating the camera and one for determining the equation of the plane of light in world coordinates. However in our case the two steps are independent while in the case of active range-sensing they are not.

### 4 A two-step calibration method

We begin by describing the structure of the calibrating object being used by our method. The choice for this structure will soon be made clear. The calibrating object is shown on Figure 2. It is a plane on which four straight lines are drawn,  $D_1, D_2, D_3, D_4$ . The first three are mutually parallel and  $D_4$  makes an angle with the direction of the first three ones. The equations of these lines are known in the calibrating (world) coordinate frame that is defined as follows: the origin of this frame is not constrained to a particular position. The  $X$ -axis is parallel to  $D_1$  and the  $Y$ -axis is perpendicular to  $D_1$ . The  $Z$ -axis is perpendicular to the  $X - Y$  plane. In this frame and in the plane  $Z = 0$ , the equations of the four lines can be written as:

$$Y = 0 \quad (D_1)$$

$$\begin{aligned} Y &= \alpha & (D_2) \\ Y &= \beta & (D_3) \\ Y &= \gamma X + \delta & (D_4) \end{aligned}$$

where the parameters  $\alpha, \beta, \gamma,$  and  $\delta$  are fixed and they determine the structure of the calibrating object.

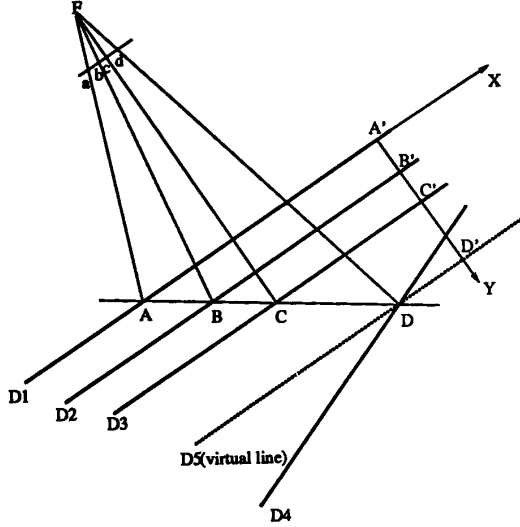


Figure 2: The structure of the calibrating object being viewed by the linear camera

#### 4.1 Estimating $n_1, n_2, n_3, n_4, n_5$

When a linear camera “looks” at this object, it “sees” four image points  $a, b, c, d$  which are the projections of  $A, B, C, D$ , the intersections of the viewing plane with  $D_1, D_2, D_3, D_4$ , e.g., Figure 2. Notice that the exact locations in space of  $A, B, C, D$  are not known. Instead, due to the structure of the calibrating object and the particular choice of the calibrating frame, the  $Y$ - and  $Z$ -coordinates of  $A, B, C$  are known independently of the position and orientation of the viewing plane.

Therefore one can establish the following correspondences:  $\{Y_A, Z_A, u_a\}$ ,  $\{Y_B, Z_B, u_b\}$ , and  $\{Y_C, Z_C, u_c\}$ . By moving the calibration object in the  $Y$ -direction and/or in the  $Z$ -direction with known increments, one may establish three correspondences at each new position of the calibrating object. Each correspondence  $i$  verify equation (4) which can be written as:

$$Y_i n_1 + Z_i n_2 + n_3 - u_i Y_i n_4 - u_i Z_i n_5 = u_i \quad (5)$$

For  $n$  such correspondences,  $n \geq 5$ , we obtain:

$$\begin{pmatrix} Y_1 & Z_1 & 1 & -u_1 Y_1 & -u_1 Z_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ Y_n & Z_n & 1 & -u_n Y_n & -u_n Z_n \end{pmatrix} \begin{pmatrix} n_1 \\ n_2 \\ n_3 \\ n_4 \\ n_5 \end{pmatrix} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \quad (6)$$

This is an overconstrained system of  $n$  linear equations in five unknowns of the form:  $AX = B$ . By pre-multiplying with the transpose of  $A$  we obtain a linear system with five equations and five unknowns:  $A^t A X = A^t B$  for which robust numerical solutions, such as Gauss-Jordan elimination, are available [4].

#### 4.2 Estimating $p, q, r$

In order to estimate  $p, q,$  and  $r$ , sets of 3-D coordinates of points belonging to the viewing plane must be available. Obviously a linear camera cannot provide point-to-point correspondences in the general case. The structure of the calibrating object has been purposively designed to be able to provide, in conjunction with the camera, 3-D point coordinates that belong to the viewing plane.

We recall the definition of the cross-ratio of four collinear points:

$$\begin{aligned} \text{cross-ratio}_{A,B,C,D} &= (A, B, C, D) \\ &= (CA/CB)/(DA/DB) \end{aligned}$$

$CA$  being the algebraic distance from  $C$  to  $A$ . The fundamental property of the cross-ratio is that it is invariant by central or parallel projection. Therefore we have:

$$(A, B, C, D) = (a, b, c, d)$$

We consider a *virtual* line  $D_5$  parallel to  $D_1, D_2, D_3$  and passing through the point  $D$ . These four parallel lines  $D_1, D_2, D_3$  and  $D_5$  intersect the  $Y$ -axis in  $A', B', C',$  and  $D'$ . From the projective invariance of the cross-ratio we have:

$$(A, B, C, D) = (A', B', C', D')$$

and hence:

$$(C'A'/C'B')/(D'A'/D'B') = (ca/cb)/(da/db)$$

We conclude that: if the linear camera sees simultaneously  $D_1, D_2, D_3,$  and  $D_4$  then one is able to compute the cross-ratio  $r, r = (a, b, c, d)$  from the image

measurements and therefore determine the position of  $D'$  along the  $Y$ -axis, say  $Y = \lambda$ .<sup>1</sup>

The coordinates of  $D$  are therefore given by the intersection of the object line  $D_4$  with the virtual line  $D_5$ :

$$\begin{aligned} Y &= \gamma X + \delta \\ Y &= \lambda \\ Z &= 0 \end{aligned}$$

Moreover, the point  $D$  is guaranteed to belong to the viewing plane and its position is computed via image measurements independently of the camera parameters. By moving the calibration object in the  $Y$ -direction and/or in the  $Z$ -direction with known increments, one may determine the coordinates of one viewing-plane point at each new position of the calibrating object. Thus we obtain  $k$  points,  $k \geq 3$ , each point  $j$  verifying the equation:

$$Y_j p + Z_j q + r = X_j$$

By combining these equations we obtain:

$$\begin{pmatrix} Y_1 & Z_1 & 1 \\ \vdots & \vdots & \vdots \\ Y_k & Z_k & 1 \end{pmatrix} \begin{pmatrix} p \\ q \\ r \end{pmatrix} = \begin{pmatrix} X_1 \\ \vdots \\ X_k \end{pmatrix} \quad (7)$$

This system of equations has the same structure as equation (6) and hence the same method can be used for finding a solution.

## 5 The extrinsic parameters

The linear camera model is described by eight parameters that we just defined and estimated. We recall that this model describe a line passing through the center of projection and an image pixel, in world frame. This frame may well be viewed as the sensor frame and for a number of applications it is sufficient to have such a model. Nevertheless, in some applications of camera calibration, such as hand/eye calibration, it is necessary to calibrate the camera several times, from several viewpoints with respect to the calibration object and to determine the rigid transformation between a camera frame and the calibrating frame [7]. The parameters of this rigid transformation are also called the extrinsic camera parameters and they will be made explicit in this Section.

<sup>1</sup>It is straightforward to derive the value of  $\lambda$ :  $\lambda = \alpha\beta / (r\alpha + (1-r)\beta)$ .

First we define a camera frame, e.g., Figure 1. The origin of the camera frame is the center of projection  $F$ . Therefore the translation vector associated with the transformation is the vector from  $O$  to  $F$  having as coordinates the coordinates of  $F$  in the calibrating frame:

$$\vec{t} = \begin{pmatrix} X_F \\ Y_F \\ Z_F \end{pmatrix}$$

The axes of the camera cartesian coordinate frame are defined by the following vectors:

- the direction of an arbitrarily chosen viewing line,  $\vec{l}$ ;
- the normal to the viewing plane,  $\vec{m}$ , and
- the cross-product of these vectors,  $\vec{n}$ .

Without loss of generality we suppose that the camera has 1024 pixels,  $u_1$  through  $u_{1024}$ . The viewing line passing through the central pixel  $u_{512}$  is given by the following equations (see equations (3) and (4)):

$$\begin{aligned} (n_1 - n_4 u_{512})Y + (n_2 - n_5 u_{512})Z + n_3 - u_{512} &= 0 \\ -X + pY + qZ + r &= 0 \end{aligned}$$

The direction of this line,  $\vec{l}$  is given by the cross-product of the normals to these two planes. These two normals are:

$$\begin{aligned} \vec{N}_1 &= (0 \quad n_1 - n_4 u_{512} \quad n_2 - n_5 u_{512})^t \\ \vec{N}_2 &= (-1 \quad p \quad q)^t \end{aligned}$$

The axes of the camera frame have the following unit direction vectors:

$$\begin{aligned} \vec{l} &= \frac{\vec{N}_1 \times \vec{N}_2}{\|\vec{N}_1 \times \vec{N}_2\|} \\ \vec{m} &= \frac{\vec{N}_2}{\|\vec{N}_2\|} \\ \vec{n} &= \vec{l} \times \vec{m} \end{aligned}$$

These three vectors determine the rotation matrix associated with the rigid transformation:

$$R = (\vec{l} \vec{m} \vec{n})$$

To determine the translation one may easily notice that all the viewing lines pass through the center of projection, in particular the viewing lines associated

with the first and last pixels (or two any other pixels). The coordinates of  $F$  are given by:

$$\begin{aligned}(n_1 - n_4 u_1)Y_F + (n_2 - n_5 u_1)Z_F + n_3 - u_1 &= 0 \\(n_1 - n_4 u_{1024})Y_F + (n_2 - n_5 u_{1024})Z_F + n_3 - u_{1024} &= 0 \\-X_F + pY_F + qZ_F + r &= 0\end{aligned}$$

It is worthwhile to notice that the extrinsic parameters embedded by the rotation matrix  $R$  and the translation vector  $\vec{t}$  have been determined geometrically without making explicit the intrinsic camera model.

## 6 Results and conclusion

Cybernetix currently develops sensor-based robotic systems for performing various industrial and service tasks. Among these systems, inspection of 3-D parts for accurate machining requires that both the camera and the machining tool are precisely positioned with respect to the 3-D part.

In the setup conceived and developed by Cybernetix in collaboration with LIFIA, the linear camera and a machining burr are mounted onto a robot arm, Figure 3 and Figure 4. The arm has the task of positioning the camera at a predetermined distance from an object and with a predetermined orientation such that the object can be inspected. If any seams are detected onto the object's surface, then the burr mounted onto the robot performs fine seam smoothing. In order to perform these tasks, the rigid transformations from the camera frame to the robot frame and from the camera frame to the tool frame must be known. This is done using a standard hand/eye calibration method. Nevertheless, prior to hand/eye calibration, the camera itself must be calibrated and its extrinsic parameters extracted [7].

Using the method described herein and for one calibration we used 150 measurements for estimating the  $n_i$ 's and 50 measurements for estimating  $p$ ,  $q$ , and  $r$ . An important merit of a calibration procedure is its accuracy. Table 1 and Table 2 summarize the results of one calibration. Together with the values obtained for the sought parameters the covariance matrices are provided as well.

The only parameter which is not accurately estimated is  $r$ , the intersection of the viewing plane with the  $X$ -axis. The reason for this problem is that the set

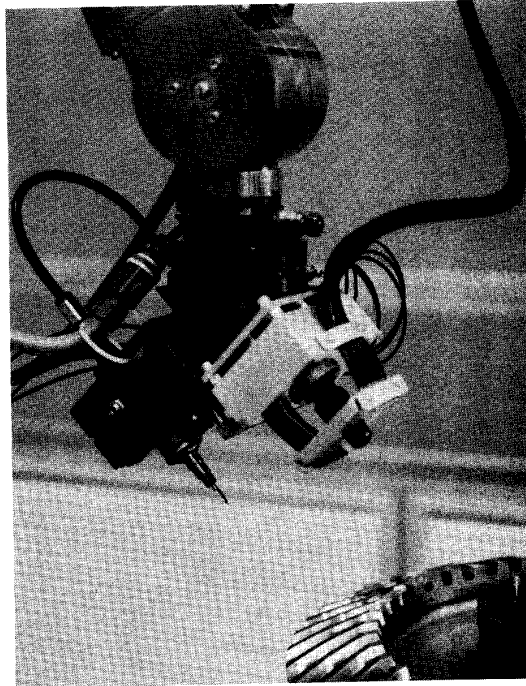


Figure 3: Both the linear camera and the machining tool are mounted onto the robot's sixth axis.

of 3-D points that were used for estimating the viewing plane equation were not evenly distributed in this plane but rather concentrated. One way to overcome this problem is to consider a calibrating object that has more than one oblique line. Nevertheless the current precision of the inspection system is of 1/50mm.

## Appendix: Why eight parameters?

The linear-camera model describes the central projection of a 3-D point onto a straight line (the linear set of pixels). The 3-D point being in general described in a world frame that is different from the camera frame, a frame change must be performed prior to the projection. Six parameters are necessary to describe the frame change transformation, three for rotation and three for translation.

The camera frame is chosen such that the linear image is perpendicular to one of the axes (the optical axis) and the origin is at the center of projection. Two parameters are necessary to describe this central projection (four in the case of a 2-D image): the pixel coordinate of the intersection of the image with the optical axis and the ratio of the focal length with the

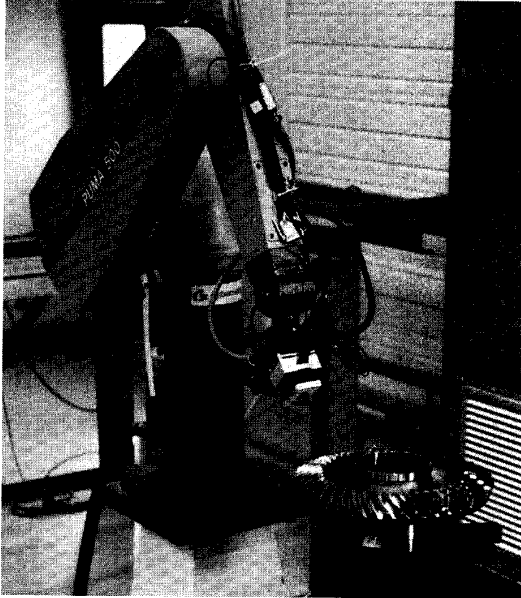


Figure 4: A global view of the inspection and fine-machining system.

pixel size.

It is straightforward to see that the variation of the value of one of these 8 parameters modifies the camera model and that this modification cannot be compensated by the variation of any of the remaining parameters. Hence the model is described by 8 independent parameters.

## References

[1] R. C. Bolles, J. H. Kremers, and R. A. Cain.

$n_1$	$n_2$	$n_3$	$n_4$	$n_5$
46.76	7.47	130.62	0.0008	0.0122
Covariance matrix				
0.0083	0.0006	0.039	0.0	0.0
0.0006	0.0136	0.001	0.0	0.0
0.039	0.001	0.197	0.0	0.0
0.0	0.0	0.0	0.0	0.0
0.0	0.0	0.0	0.0	0.0

Table 1: Nominal values and covariance matrix for the first step of the calibration method

$p$	$q$	$r$
-0.434	-0.023	18.836
Covariance matrix		
0.011	0.0008	-0.216
0.0008	0.0187	-0.0115
-0.216	-0.0115	4.269

Table 2: Nominal values and covariance matrix for the second step of the calibration method

A Simple Sensor to Gather Three-Dimensional Data. Technical Report 249, SRI International, July 1981.

[2] O. D. Faugeras and G. Toscani. The Calibration Problem for Stereo. In *Proc. Computer Vision and Pattern Recognition*, pages 15–20, Miami Beach, Florida, USA, June 1986.

[3] R. Mohr and L. Morin. Relative positioning from geometric invariants. In *Proceedings Computer Vision and Pattern Recognition Conference*, pages 139–144, Lahaina, Maui, Hawaii, June 1991.

[4] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Wetterling. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, 1988.

[5] P. Puget and Th. Skordas. An Optimal Solution for Mobile Camera Calibration. In O. Faugeras, editor, *Computer Vision – ECCV 90, Proceedings First European Conference on Computer Vision, Antibes, France*, pages 187–198. Springer Verlag, April 1990.

[6] R. Y. Tsai. A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses. *IEEE Journal of Robotics and Automation*, RA-3(4):323–344, August 1987.

[7] R.Y. Tsai and R.K. Lenz. Real Time Versatile Robotics Hand/Eye Calibration using 3D Machine Vision. In *IEEE International Conference on Robotics and Automation*, pages 554–561, Philadelphia, Penn, USA, April 1988.