International Journal of Computer Vision, 4, 127–140 (1990)

Using Vanishing Points for Camera Calibration

B. CAPRILE AND V. TORRE
Dipartimento di Fisica dell'Universitá di Genova, Via Dodecaneso 33, 16146 Genova, Italy

Abstract
In this article a new method for the calibration of a vision system which consists of two (or more) cameras is presented. The proposed method, which uses simple properties of vanishing points, is divided into two steps. In the first step, the intrinsic parameters of each camera, that is, the focal length and the location of the intersection between the optical axis and the image plane, are recovered from a single image of a cube. In the second step, the extrinsic parameters of a pair of cameras, that is, the rotation matrix and the translation vector which describe the rigid motion between the coordinate systems fixed in the two cameras, are estimated from an image stereo pair of a suitable planar pattern. Firstly, by matching the corresponding vanishing points in the two images the rotation matrix can be computed, then the translation vector is estimated by means of a simple triangulation. The robustness of the method against noise is discussed, and the conditions for optimal estimation of the rotation matrix are derived. Extensive experimentation shows that the precision that can be achieved with the proposed method is sufficient to efficiently perform machine vision tasks that require camera calibration, like depth from stereo and motion from image sequence.

1 Introduction

An essential step for any stereo vision system [1, 2, 8] is the evaluation of the geometry of the cameras, that is, of the geometrical properties of the imaging process within a single camera (intrinsic parameters), of the rotation matrix **R**, and of the translation vector **T** between pairs of cameras (extrinsic parameters). The accuracy of the measurements of these quantities is essential for a full exploitation of the epipolar constraint [1, 3] and for an accurate reconstruction of 3D data. Tsai [12, 13], Faugeras and Toscani [4] have already presented very good techniques to evaluate both the intrinsic and extrinsic parameters of a stereo system, using optical calibration without using direct mechanical measurements. Optical calibration allows the reconstruction of camera geometry by simply viewing a suitable calibrating pattern without using direct measurements, which are impractical in many real environments.

In this article we present a new and simple technique for optical calibration. By using properties of vanishing points [7, 9, 10], it allows the calculation of camera geometry by using only a single view of a calibrating pattern. The proposed technique is robust against noise and very transparent. Moreover, it can be used by a robot to navigate in an environment where vanishing points can be detected. A preliminary report of this work has already been presented [14].

2 Properties of Vanishing Points

This section establishes notations and discusses a few basic properties of vanishing points which are very useful for the purpose of camera calibration.

In what follows, let x, y, z be an orthogonal system of coordinates (whose values are meant to be in *mm*) associated with a viewing camera, such that the origin of the system coincides with the center of projection (that is, the lens center) and the z-axis with the optical axis. The image plane is defined by the equation $z = f$ where f is the focal length of the imaging device. Thus, the coordinates X, Y of the point q_p on the image plane, perspective projection of a point Q in space, expressed in pixels, are

$$\begin{cases} x = -\dfrac{f}{k_1}(Y_1 - b) \\ y = \dfrac{f}{k_2}(X - a) \end{cases} \qquad k_1, k_2 > 0 \quad (1)$$

where k_1 and k_2 are the focal lengths of the imaging device in *X-pixel* and *Y-pixel* units respectively, f is (in *mm*), and a and b are the pixel coordinates of the intersection of the optical axis with the image plane. It is evident that equation (1) can be used effectively as a relationship between the pixel-coordinates X, Y and the coordinates x, y only if $k_1/k_2, f/k_1$, and a and b are known. Let us call these quantities *intrinsic parameters*. A simple procedure for the estimate of these quantities from one image of a suitable pattern is described in section 6.1.

Let us now briefly review the concept of vanishing point and some useful properties of perspective projection [7, 9, 10]. Let us consider a straight line S in R^3 through the point (a_x, a_y, a_z), parameterized as follows:

$$\begin{cases} x = a_x + tn_x \\ y = a_y + tn_y \\ z = a_z + tn_z \end{cases} \tag{2}$$

where $\mathbf{n} = (n_x, n_y, n_z)$ is the unit vector giving the direction of S and t is a parameter.

The *vanishing point* of the straight line S is the point $\mathbf{v}_S = (x_\infty, y_\infty, z_\infty)$ on the image plane, where

$$\begin{cases} x_\infty = \lim\limits_{t\to\infty} f\dfrac{a_x + tn_x}{a_z + tn_z} = f\dfrac{n_x}{n_z} \\ y_\infty = \lim\limits_{t\to\infty} f\dfrac{a_y + tn_y}{a_z + tn_z} = f\dfrac{n_y}{n_z} \\ z_\infty = f \end{cases} \tag{3}$$

It is now useful to introduce the following definition:

DEFINITION. *Let $\mathbf{q} \equiv (x, y, f)$ be a point lying on the image plane and let $\mathbf{v}_q \equiv (x/h, y/h, f/h)$ be the unit vector associated with $\mathbf{q}$, where $h = (x^2 + y^2 + f^2)^{1/2}$. If we consider a straight line in space S and its vanishing point $\mathbf{v}_S$, it follows that the unit vector associated with $\mathbf{v}_S$ is equal to $\mathbf{n}_S$, which is the unit vector giving the direction of S (if both are assumed to point toward the image plane).*

Let us now list three properties of vanishing points which will be used in the sequel to determine the calibration parameters. The first two properties are useful for the determination of the *extrinsic* parameters—that is, the rotation matrix and the translation vector which describe the transformation between the orthogonal systems of coordinates associated with a pair of cameras.

PROPERTY 1. *Let $A \equiv \{S_i\}$ be a set of straight lines (not parallel to the image plane). The set $\alpha \equiv \{\mathbf{v}_{S_i}\}$ of the vanishing points associated with the elements of A lies on a straight line if and only if the straight lines in A are parallel to the same plane Π_A.*

Proof. If the vanishing points lie on a straight line l, it is evident that the straight lines in A are parallel to the plane Π_A which passes through the origin, center of projection, and the line l. Conversely, if straight lines in A are parallel to the same plane Π_A, and Π_α is the plane parallel to Π_A and passing through the origin, it is evident that the plane Π_α contains the vanishing points associated with each straight line in A. Thus the intersection of the image plane and Π_α, which is a straight line, contains all the vanishing points associated with the elements of A.

The next property shows how to exploit orthogonality of straight lines:

PROPERTY 2. *Let Q, R, and S be three mutually orthogonal straight lines in space and $\mathbf{v}_Q \equiv (x_Q, y_Q, f)$, $\mathbf{v}_R \equiv (x_R, y_R, f)$, $\mathbf{v}_S \equiv (x_S, y_S, f)$ the three vanishing points associated with them. If we know the coordinates of one of these points, say $\mathbf{v}_Q$, and the direction of a straight line on the image plane which passes through a second one, say $\mathbf{v}_R$, we can compute the coordinates of both $\mathbf{v}_R$ and $\mathbf{v}_S$.*

Proof. Let $\mathbf{l}$, $\mathbf{m}$, $\mathbf{n}$ be the unit vectors associated respectively with $\mathbf{v}_Q$, $\mathbf{v}_R$, $\mathbf{v}_S$. Because Q, R, and S are mutually orthogonal we have

$$\begin{cases} \mathbf{l} \cdot \mathbf{m} = 0 \\ \mathbf{l} \cdot \mathbf{n} = 0 \\ \mathbf{m} \cdot \mathbf{n} = 0 \end{cases} \tag{4}$$

so that we obtain the three equations

$$\begin{cases} x_Q x_R + y_Q y_R + f^2 = 0 \\ x_Q x_S + y_Q y_S + f^2 = 0 \\ x_R x_S + y_R y_S + f^2 = 0 \end{cases} \tag{5}$$

Then, property 2 follows easily from the hypothesis and the three equations (5).

In the next section the location of the intersection of the optical axis with the image-plane, or one of the intrinsic parameters, will be estimated by means of the following property.

PROPERTY 3. *Let Q, R, S be three mutually orthogonal straight lines in space, and let* $v_Q \equiv (x_Q, y_Q, f)$, $v_R \equiv (x_R, y_R, f)$, $v_S \equiv (x_S, y_S, f)$ *be the three vanishing points associated with them. The orthocenter of the triangle with vertexes in the three vanishing points is the intersection of the optical axis and the image plane.*

Proof. The equation of the straight line T_Q passing through v_Q and orthogonal to the segment joining v_R and v_S is

$$y - y_Q = \frac{x_S - x_R}{y_R - y_S}(x - x_Q) \tag{6}$$

From equations (5) and (1), it follows immediately that T_Q passes through $X = a$, $Y = b$. By applying the same argument to the other vanishing points property 3 follows easily.

3 Extrinsic Parameters: Rotation

A rigid displacement of a camera, or, equivalently, the relative position and orientation of a pair of cameras, is characterized by a 3×3 orthogonal matrix $\mathbf{R}$ and by a vector $\mathbf{T}$, representing respectively the rotation between the "old" and the "new" system of coordinates and the translation of the origin.

The key observation is that, by definition, a set of parallel straight lines have the same vanishing point. Thus, vanishing points (and their associated unit vectors) "do not change their coordinates" for a rigid translation of the viewing system. From this, we have that changes in the location of a vanishing point for arbitrary rigid motion of a camera, or arbitrary relative geometry between a pair of cameras, are due to the relative component of rotation. Therefore, from elementary linear algebra, we have that the coordinate transformation of a unit vector $\mathbf{v}$ which is associated with a vanishing point, for an arbitrary rigid motion between the systems of coordinates, can be expressed by

$$\mathbf{v}' = \mathbf{R}\mathbf{v} \tag{7}$$

Let us assume that the coordinates of three vanishing points not lying on the same straight lines are known. If we consider the 3×3 matrix $\mathbf{V}$ whose columns are the three unit vectors $\mathbf{v}_i$, $i = 1, 2, 3$, associated with the three vanishing points, then we can write

$$\mathbf{V}' = \mathbf{R}\mathbf{V} \tag{8}$$

Equation (8) is a system of equations for the elements of $\mathbf{R}$ that can be solved if $\mathbf{V}$ is invertible. It is obvious that

$$| \det \mathbf{V} | \equiv \mathbf{v}_1 \cdot (\mathbf{v}_2 \wedge \mathbf{v}_3) \neq 0 \tag{9}$$

if the three unit vectors are not coplanar, and we know from property 1 (see section 2) that vanishing points of noncoplanar directions are not collinear. The problem is now reduced to finding the coordinates of three noncollinear vanishing points in two images. In the case of three vanishing points associated with three orthogonal straight lines, the matrix $\mathbf{V}$ is orthonormal and we have

$$\mathbf{R} = \mathbf{V}'\mathbf{V}^{\mathrm{T}} \tag{10}$$

where $\mathbf{V}^{\mathrm{T}}$ indicates the transpose of $\mathbf{V}$. It is worthwhile stressing the extreme simplicity of equation (10) for recovering the rotation, as opposed to previous solutions [4, 6, 12].

4 Extrinsic Parameters: Translation

Once the rotation matrix has been obtained, it is possible to recover the translation vector $\mathbf{T}$ from the correspondence of the projections of a line segment of known length and orientation in two different frames. The proposed procedure for the recovering of translation is a simple triangulation.

Let $\mathbf{P}$ be the vector expressing, in the camera's system of coordinates, a point P in space. If the camera moves, it changes according to

$$\mathbf{P}' = \mathbf{R}(\mathbf{P} - \mathbf{T}) \tag{11}$$

where $\mathbf{R}$ is the rotation matrix and $\mathbf{T}$ is the translation vector. Thus, once $\mathbf{R}$ is obtained by means of the method explained in section 3 and if we know $\mathbf{P}$ and $\mathbf{P}'$, from equation (11) we obtain

$$\mathbf{T} = \mathbf{P} - \mathbf{R}^{\mathrm{T}}\mathbf{P}' \tag{12}$$

Figure 1a represents a camera in a fixed position. Let us assume that the orientation and the length of the segment joining P_1 and P_2 are known. By a simple inspection of figure 1 it is evident that the unit vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ can be computed easily because p_1 and p_2 are the

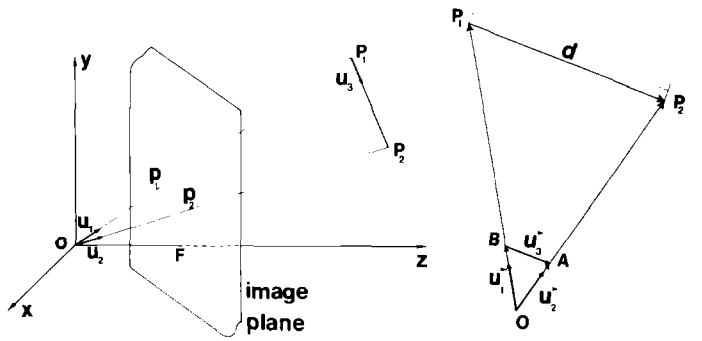

Fig. 1. Scheme of triangulation used to recover the translation vector **T**. (a) The projection p_1 and p_2 of the extrema of the segment connecting P_1 to P_2 define a pair of unit vectors $\mathbf{u}_1$, $\mathbf{u}_2$ in the camera reference system. (b) If the length and the spatial orientation of **d** is known, the coordinate of P_1 and P_2 can be obtained by a simple triangulation.

projection on the image plane of P_1 and P_2. Furthermore, $\mathbf{u}_3$ is known because the orientation of the segment joining P_1 and P_2 is known. Thus we have the following system of linear equations for the unknowns α and β (see figure 1b).

$$\alpha\mathbf{u}_1 - \mathbf{u}_3 = \beta\mathbf{u}_2 \tag{13}$$

Once α and β have been determined, the space coordinates of P_1 can be easily obtained from the similarity of the triangles OBA and OP_1P_2. If the camera moves to a new position, it is still possible to calculate the new space coordinate of P_1 by using the same method, and the translation vector **T** can be computed immediately by using equation (12).

5 Robustness of the Method

In order to evaluate the practicality of the proposed technique, it is necessary to verify its robustness against noise. Ultimately, the accuracy of the method depends on the exact localization of vanishing points and on the error propagation in the estimate of rotation and translation.

5.1 Localization of Vanishing Points

Let us now evaluate the accuracy in the localization of a vanishing point. In the image plane a vanishing point is located as the intersection of straight lines, which are the projection of parallel straight lines in the scene.

Let $\mathbf{x} \cdot \mathbf{n}_i = \gamma_i$ and $\mathbf{x} \cdot \mathbf{n}_j = \gamma_j$ be the equations of two straight lines on the image plane, where $\mathbf{x} = (x, y)$, $\mathbf{n}_i = (\cos\theta_i, \sin\theta_i)$, $\mathbf{n}_j = (\cos\theta_j, \sin\theta_j)$, and γ_i and γ_j are the distance of straight lines from the origin. The intersection $\mathbf{P} = (x_\infty, y_\infty)$ of the two straight lines is

$$x_\infty = \frac{1}{\sin(\theta_2 - \theta_1)}[\gamma_1 \sin\theta_2 - \gamma_2 \sin\theta_1]$$

$$y_\infty = \frac{1}{\sin(\theta_2 - \theta_1)}[\gamma_2 \cos\theta_1 - \gamma_1 \cos\theta_2] \tag{14}$$

and the conditioning number, c (see [11]), of this linear system is

$$c = \frac{1 + |\cos(\theta_2 - \theta_1)|}{1 - |\cos(\theta_2 - \theta_1)|} \tag{15}$$

In order to have robustness against noise the conditioning number c must be as close as possible to 1 and therefore it is desirable to have $\theta_2 - \theta_1$ close 90°. In order to increase the accuracy of the localization of the vanishing point it is useful also to locate it as the intersection of many straight lines. Given n parallel straight lines in the scene, their vanishing point $\bar{\mathbf{x}} = (x_\infty, y_\infty)$ can be found as the average of the $n(n-1)/2$ intersections of the different straight lines on the image plane,

$$\bar{\mathbf{x}} = (x_\infty, y_\infty) = \frac{2}{n(n-1)} \sum_{k=1}^{n(n-1)/2} \bar{\mathbf{x}}_k \tag{16}$$

where $\bar{\mathbf{x}}_k$ is the intersection of a pair of straight lines. It was found convenient and reliable to slightly modify equation (16), so as to give more weight to estimates of the location of vanishing points from intersections of straight lines not parallel. Therefore we used the formula

$$\bar{\mathbf{x}}_p = (x_\infty, y_p) = \sum_{k=1}^{n(n-1)/2} p_k \bar{\mathbf{x}}_k \tag{17}$$

with

$$\sum_{k=1}^{n(n-1)/2} p_k = 1$$

and p_k proportional to $\theta_i - \theta_j$.

In order to establish the overall accuracy of the localization of the vanishing point we performed computer simulations. We considered that the end points of segments on the image, used to find the intersection of straight lines, had an error of one or two pixels and we calculated the uncertainty on the localization of the vanishing points. From these computer simulations the absolute error of the localization of the vanishing point of our calibrating pattern was about 1 pixel in abscissa and 2 pixels in ordinate. This accuracy, as shown in a later section, is adequate for a reliable estimate of calibration parameters.

5.2 The Computation of Rotation from Vanishing Points

From equation (8) the rotation matrix $\mathbf{R}$ is given by

$$\mathbf{R} = \mathbf{V}' \mathbf{V}^{-1} \tag{18}$$

Standard error analysis [11] shows that the error propagation in (18) is determined by the conditioning number c of the matrix $\mathbf{V}$. The conditioning number c of $\mathbf{V}$ is

$$c = \frac{\gamma_{max}}{\gamma_{min}} \tag{19}$$

where γ_{max} and γ_{min} are, in absolute value, the largest and the smallest eigenvalue of the matrix $\mathbf{V}^T\mathbf{V}$, respectively. By indicating the columns of $\mathbf{V}$ with vectors $\mathbf{l}$, $\mathbf{m}$ and $\mathbf{n}$ then the characteristic equation of the symmetric matrix $\mathbf{V}^T\mathbf{V}$ is

$$(1 - \gamma)\{(1 - \gamma^2) - \cos^2\theta_{lm} - \cos^2\theta_{ln} - \cos^2\theta_{mn}\} = 0 \tag{20}$$

where $\cos\theta_{lm} = \mathbf{l} \cdot \mathbf{m}$, $\cos\theta_{ln} = \mathbf{l} \cdot \mathbf{n}$ and $\cos\theta_{mn} = \mathbf{m} \cdot \mathbf{n}$.

The conditioning number c is then

$$c = \frac{1 + (\cos^2\theta_{lm} + \cos^2\theta_{ln} + \cos^2\theta_{mn})^{1/2}}{1 - (\cos^2\theta_{lm} + \cos^2\theta_{ln} + \cos^2\theta_{mn})|^{1/2}} \tag{21}$$

As a consequence the conditioning number c is equal to 1 when $\mathbf{l}$, $\mathbf{m}$ and $\mathbf{n}$ are mutually orthogonal, thus giving best robustness against noise. In this case $\mathbf{V}^T\mathbf{V} = \mathbf{I}$ and there is no amplification of errors due to the localization of vanishing points. In order to see more quantitatively the effect of an error in the localization of the vanishing point on the elements of the matrix $\mathbf{R}$ we performed some computer simulations. The result of this analysis showed that an error of the localization of a vanishing point of about 2 pixels caused a relative error of less than 0.1% in the values of R.

5.3 The Computation of Translation

The estimate of translation, as described in section 4, is obtained with a simple triangulation. Its accuracy depends on the exact measure of the length of the calibrating segment and on the localization of end points on the image plane. All these measurements can have an accuracy of at least 1 mm and 1 pixel respectively. Computer simulation of the error propagation showed that in the experimental configuration of the calibrating pattern, the translation could be estimated with an absolute error of 1 or 2 mm according to the geometry used. The relative error could be less than 1%.

6 Camera Calibration

This section describes the procedures implemented for the determination of intrinsic and extrinsic parameters of a pair of cameras which were looking at the same scene.

Any stereo vision system that is able to obtain quantitative depth information about a scene requires the knowledge of the transformation linking the two systems of coordinates defined by the two cameras [5]. It is important to distinguish between parameters describing the relative spatial position of the cameras, or extrinsic parameters, and quantities characterizing the cameras themselves, or intrinsic parameters. Elementary considerations of experimental physics suggest the use of as many direct measurements of the quantities of interest as possible. We have devised an off-line procedure to measure first the intrinsic parameters and then an on-line procedure to recover the extrinsic parameters.

6.1 Estimation of Intrinsic Parameters

The intrinsic parameters are,

1. the focal length of the device expressed in *X*-pixel units k_2,
2. the ratio k_1/k_2 between the values of the focal length of the device expressed in *Y*-pixel and in *X*-pixel units,
3. the pixel coordinates of the intersection between the optical axis and the image plane, a and b.

The value of k_1/k_2 results from geometrical properties of the photosensors matrix and from the analog-to-digital conversion performed by the acquisition hardware. Several methods to evaluate k_1/k_2 have been proposed (see [6] for example). Here, it is thought that a direct and manual measurement is the most reliable. Figure 2 reproduces images obtained by closely opposing a microscope ocular with a square grid against the camera objective. In this way it is possible to project on the photosensor an almost perfectly rectangular pattern, so that the parameter k_1/k_2 can be immediately recovered simply by reading the pixels; and any possible geometrical distortions can be checked.

Using this procedure, for PULNIX TM46 cameras connected with an Imaging Technology A/D board (FG 100), we obtain a direct measure for $k_1/k_2 = 1.47$. The accuracy of this measurement is high (about 99%), because it is just a matter of counting pixels on the screen.

In exploiting properties of vanishing points, it is possible to estimate the values of the parameters a, b, and k_1. Figure 3 represents an image of a cube as it appears on the screen of the acquisition system. On three sides of the cube parallel segments were drawn, in order to have three mutually orthogonal families of parallel lines. Given one image of the cube, it is possible to locate on the image plane the three vanishing points and then, by using property 3, to recover the values of intrinsic parameters.

Because of orthogonality, the three unit vectors $\mathbf{l}$, $\mathbf{m}$, $\mathbf{n}$ associated with the three vanishing points satisfy the relations of equation (4) so that, if (X_i, Y_i), $i = 1, 2, 3$, are the pixel coordinates of these vanishing points, from equations (4) and (1) after some rearrangements we obtain

$$\begin{cases} (X_3 - X_1)a + k^2(Y_3 - Y_1)b + (X_1 - X_3)X_2 \\ \qquad + k^2(Y_1 - Y_3)Y_2 = 0 \\ (X_3 - X_2)a + k^2(Y_3 - Y_2)b + (X_2 - X_3)X_1 \\ \qquad + k^2(Y_2 - Y_3)Y_1 = 0 \end{cases} \tag{22}$$

where $k = k_1/k_2$. Once a and b are obtained from equations (4) and (22), it is straightforward to obtain k_1. The accuracy in the localization of the three vanishing points in the image of the calibrating cube increases if the cube is positioned so that its visible faces form with the image plane an angle as close to 45 degrees as possible.

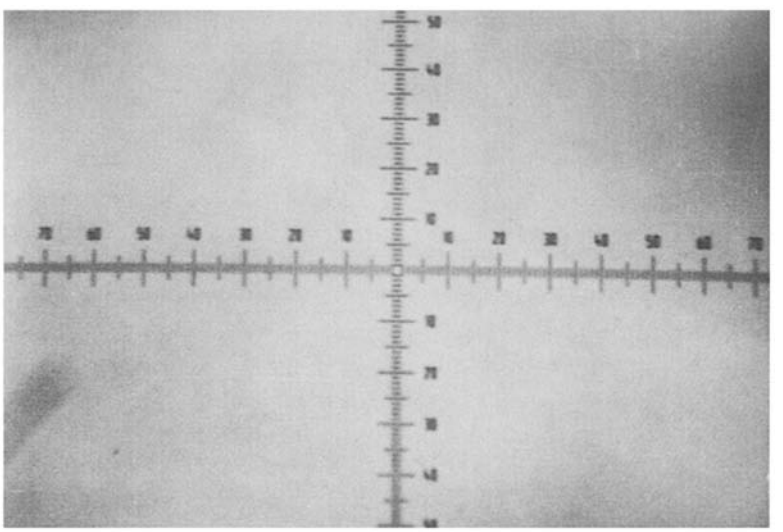

Fig. 2. Image of the calibration grid. A square calibration grid has been used to compute the ratio between the focal length in *Y*-pixel and in *X*-pixel units. Using Pulnix TM46 cameras and an Imaging Technology FG100 board we have obtain $k_1/k_2 = 1.47$.

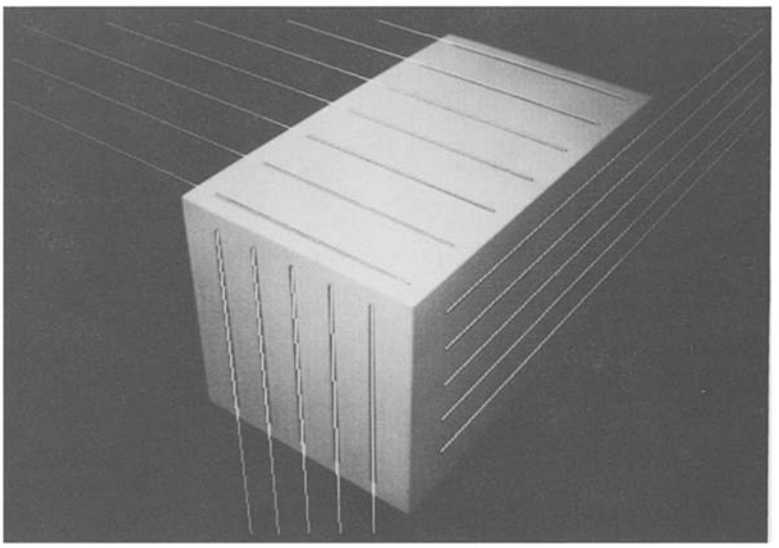

Fig. 3. The aluminum block used to calibrate three intrinsic parameters of each camera. Linear interpolation of the straight segments painted on its sides gives the location of the vanishing points associated with three mutually orthogonal directions.

The measured value of a lies, for PULNIX TM46 cameras and our acquisition system, in the range 265 $\div$ 268 pixel, while values of b lie in the range 252 $\div$ 254 pixel. The values of k_1 lie in the ranges 988 $\div$ 993 pixel and 2015 $\div$ 2021 pixel, by using objectives of the nominal focal length of 8 mm and 16 mm respectively.

6.2 Estimation of Extrinsic Parameters

When the intrinsic parameters of each camera have been obtained as described in the previous paragraph, it is possible to use the technique outlined in sections 3 and 4 to recover the rotation and the translation between a pair of cameras. Our experimental set-up was composed of a pair or triple of cameras mounted on a rail of an optical bench (see figure 4).

The cameras were looking at a table, on which a planar calibrating pattern was displayed. This pattern was composed of five parallel straight segments 35.0 cm long and distanced 4.0 cm from each other. At the "bottom" of the five straight segments there was another segment 16.0 cm long, and orthogonal to the others. Two images of the calibrating pattern (one for each camera) were analyzed in the following way:

1. The image was convolved with the Laplacian-of-Gaussian filter with $\sigma = 1.0$ pixel and zero-crossings were extracted.
2. Zero-crossing contours of the six segments were interpolated with straight lines using mean square techniques.
3. The first vanishing point was located at the intersection of the family of parallel straight lines, and the location of the second vanishing point was obtained by using property 2.
4. Once the two unit vectors associated with these two vanishing points were recovered, the third one was obtained by a simple cross-product.

The whole procedure is fully implemented on a SUN 3 workstation. Having obtained, for each camera, three unit vectors associated with mutually orthogonal straight lines it is possible to recover the rotation between the cameras as explained in section 3. Because the distance between the five parallel segments is known, the translation **T** between the cameras can also be obtained as described in section 4.

7 Experimental Results

The technique for camera calibration described in the previous section has been tested in three ways, which represent critical tests for stereo vision:

1. Comparison of the values of **R** and **T** as obtained by our calibration procedure with values of **R** and **T** measured by direct techniques.

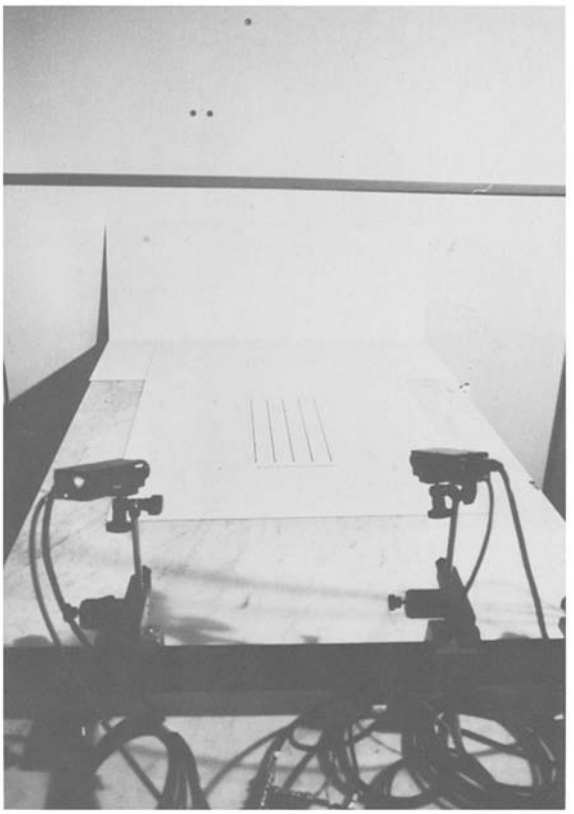

Fig. 4. The calibrating set-up. A pair of Pulnix TM46 cameras is placed in front of a planar pattern containing five parallel straight segments 35 cm long, 4 cm distant, and another segment 16 cm long orthogonal to the others. Linear interpolation of the straight segments gives the location of vanishing points.

2. Control of the viability for stereo matching of the epipolar constraint based on our calibration data.
3. Estimation of the length of a known segment by a stereo algorithm which uses the obtained calibration data.

We consider these tests to be necessary requirements for any optical calibration technique.

7.1 First Test

Because of the poor accessibility of the chip itself, the exact location of the sensor array in a camera cannot be obtained with an error smaller than 1 mm. As a consequence, a direct measure (with the usual measuring devices, such as calipers or rulers) of **R** and **T** between any two cameras is affected by an absolute error of about 2×10^{-3} and 2 mm respectively.

This error can be reduced when we slide, by a known displacement, a camera mounted on a rail of an optical bench, thus obtaining two images obtained with the same camera. Figure 5 shows a sequence of four images of the calibrating pattern viewed by the same camera, which was displaced without any intended rotation. Using a pair of these images we expect to obtain, from our algorithm, a rotation matrix very close to the identity. We can also verify the coincidence between the measured and the computed length of the translation.

In table 1 we reproduce the results obtained in 20 experiments for each distance.

Table 1. In each row of the table are reported, from left to right, the value of the pure translation of the camera directly measured on the rail of the optical bench, the correspondent mean value obtained by using our algorithm, and the maximum error on the values of the rotation matrix elements that we have met.

Direct Measure of Translation	Computed Translation	$\mathbf{R}_{ij}$ Maximum Error
20 ± 1 mm	20.5 ± 0.4 mm	$3.4\ 10^{-3}$
30 ± 1 mm	30.6 ± 0.4 mm	$4.1\ 10^{-3}$
40 ± 1 mm	40.4 ± 0.5 mm	$2.9\ 10^{-3}$
50 ± 1 mm	50.2 ± 0.3 mm	$0.9\ 10^{-3}$
60 ± 1 mm	59.6 ± 0.5 mm	$5.4\ 10^{-3}$

The computed rotation matrix is close to the identity with an error less than 1%. In experiments in which two cameras were mounted on the rail of the optical bench, separated by a distance between 15 and 45 cm with a vergence angle between 15° and 30°, we obtained the results reported in table 2. We also tried to recover the extrinsic parameters when the optical axes of viewing cameras were not coplanar. Under these conditions the calibrating pattern cannot always be simultaneously viewed by the two cameras and a good calibration with the proposed technique is difficult. However, when the two cameras were able to view a large portion of the calibrating pattern, the calibration data were in good agreement with the measured values. From these results we see that our procedure seems satisfactory.

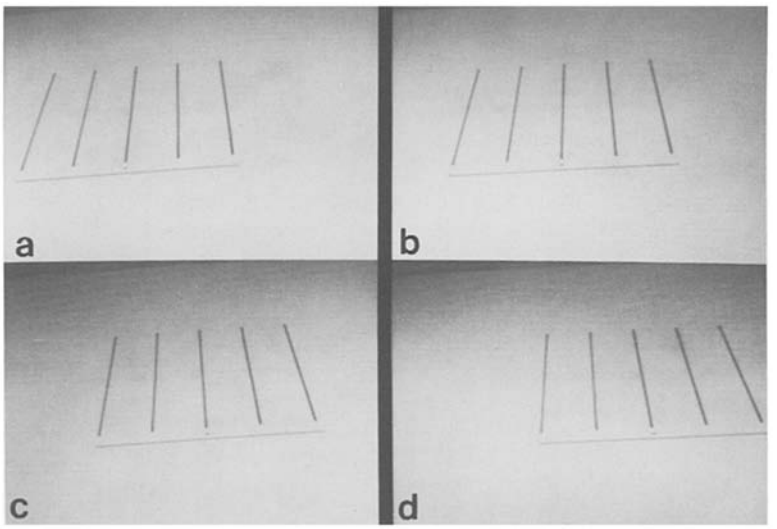

Fig. 5. Sequence of the calibrating pattern. Frame (a) is the starting point of the sequence, while frames (b), (c), (d) correspond to a translation of the camera of 25, 45, 65 cm respectively.

Table 2. In each row are reported the value of translation between cameras directly measured, the computed translation, the vergence angle between cameras, and the maximum error on the values of the rotation matrix elements. The optical axes of cameras were approximately coplanar and parallel to the standing table.

Direct Measure of Translation	Computed Translation	Vergence Angle	$\mathbf{R}_{ij}$ Maximum Error
13.6 ± 0.2 cm	13.9 ± 0.3 cm	20.7° ± 0.2	$7\ 10^{-3}$
15.2 ± 0.2 cm	15.0 ± 0.3 cm	22.1° ± 0.2	$5\ 10^{-3}$
22.1 ± 0.2 cm	21.8 ± 0.3 cm	24.0° ± 0.2	$3\ 10^{-3}$
24.0 ± 0.2 cm	24.2 ± 0.3 cm	29.0° ± 0.2	$8\ 10^{-3}$
25.8 ± 0.2 cm	25.8 ± 0.3 cm	25.3° ± 0.2	$4.2\ 10^{-3}$
29.3 ± 0.2 cm	29.3 ± 0.3 cm	31.8° ± 0.2	$2.8\ 10^{-3}$
32.5 ± 0.2 cm	32.7 ± 0.3 cm	28.3° ± 0.2	$6\ 10^{-3}$
38.4 ± 0.2 cm	38.7 ± 0.3 cm	34.3° ± 0.2	$2\ 10^{-3}$
40.3 ± 0.2 cm	41.0 ± 0.3 cm	30.3° ± 0.2	$5\ 10^{-3}$
44.7 ± 0.2 cm	44.5 ± 0.3 cm	35.2° ± 0.2	$1.5\ 10^{-3}$

Other series of tests on the accuracy of the estimates of the calibration parameters were performed by mounting one of the two cameras on a platform which could be moved by a rotary and a linear stepping motor (Compumotor). With this experimental arrangement the relative position of the two cameras could be controlled with a precision of a few microns and a tenth of an arc. The obtained results were similar to those reported in table 1 and table 2, suggesting that the proposed procedure is able to recover the extrinsic parameters with an error of less than 1%.

7.2 Second Test

The second test aims to establish whether the calibration parameters are precise enough to provide a good exploitation of the epipolar constraint. The epipolar constraint is best exploited when the two images are rectified, example, when conjugated epipolar lines become collinear lines. In figure 6a we reproduce a pair of images of objects on a table. When the two images are reprojected on a plane through the optical centers of the two cameras, conjugate points (example, the perspective projections on the two screens of the same physical point) on the two images lie on parallel lines which are the i rectified conjugate epipolar lines. The accuracy of this transformation depends on the accuracy of the calibration data. In figure 6b we show the same pair of images after the epipolar transformation. In order to facilitate a visual evaluation, pairs of conjugated epipolar lines have been traced in black.

As shown in figure 6, conjugate points lie, after the epipolar transformation, on collinear epipolar lines. Figure 7 and figure 8 show other pairs of stereo images before and after rectification. An extensive experimen-

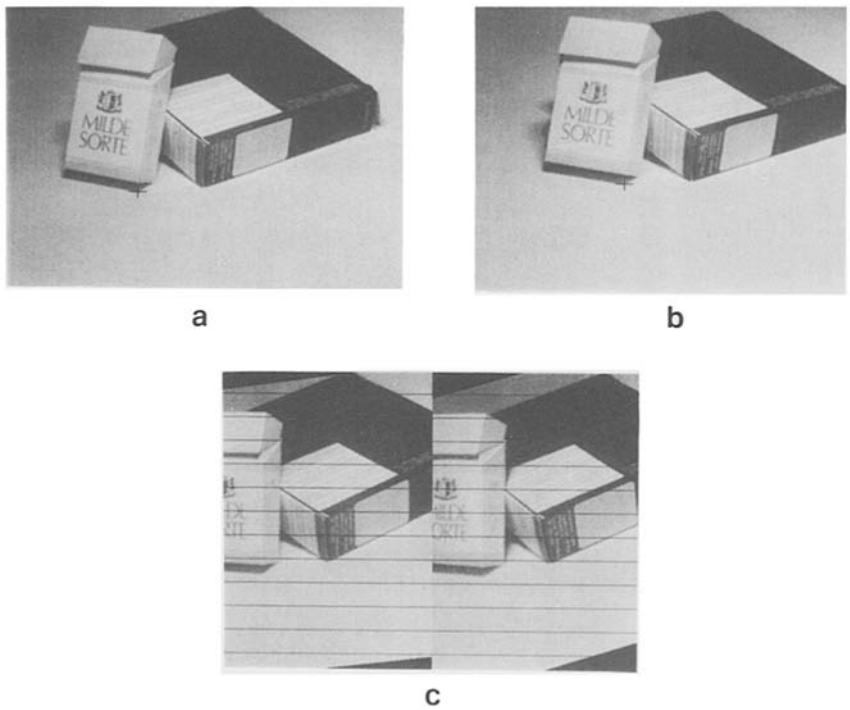

Fig. 6. (a) Pair of stereo images. (b) The images shown in b after the epipolar transformation, which uses the calibration parameters obtained by the proposed procedure. The epipolar transformation was obtained by reprojecting the two images on a plane through the optical centers of the two cameras.

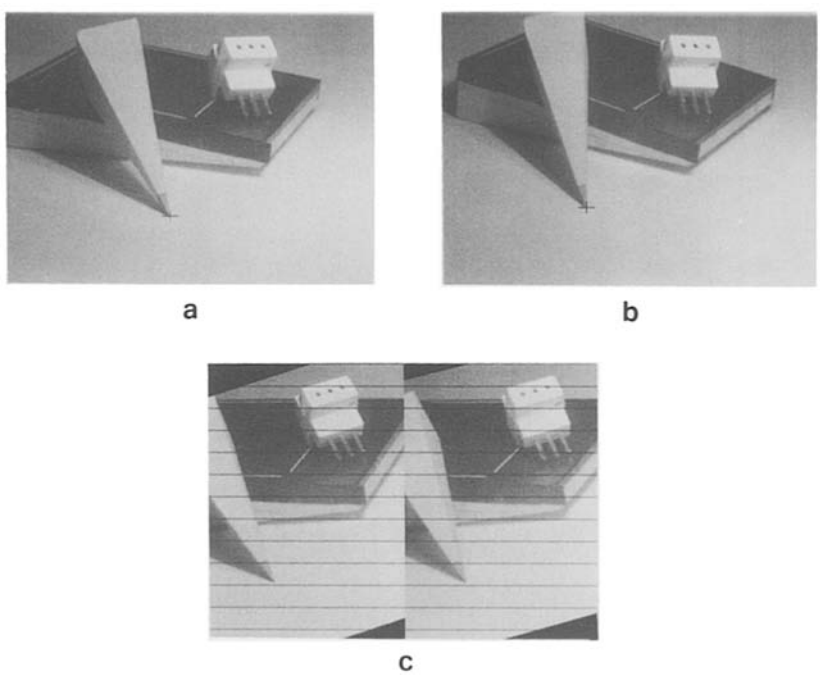

Fig. 7. Same as in figure 6.

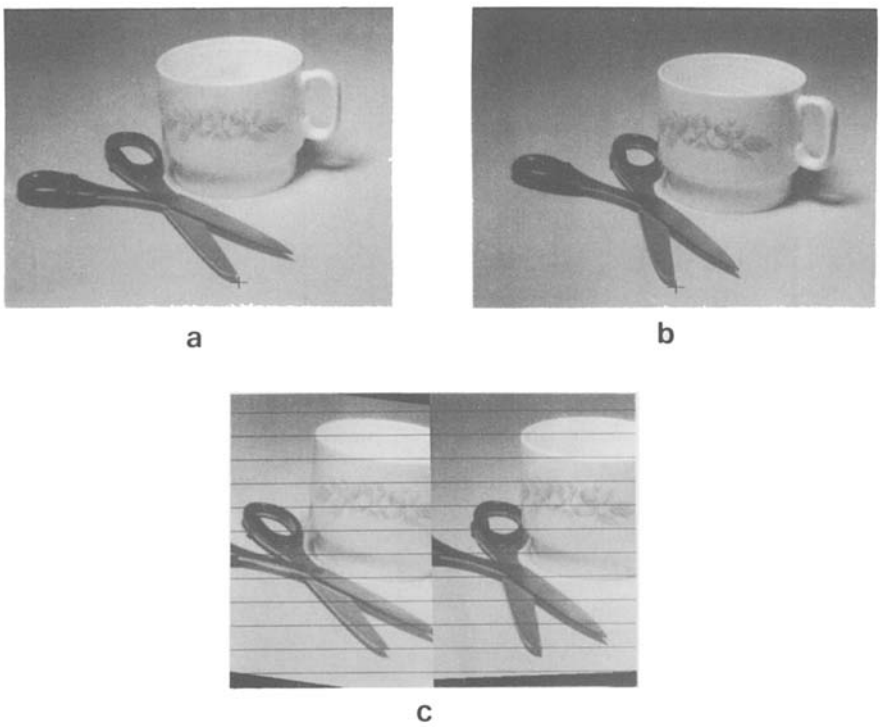

Fig. 8. Same as in figure 6.

tation on different pairs of stereo images has shown that the calibration data, obtained with the proposed procedure, are adequate to fully exploit the epipolar constraint.

7.3 Third Test

The final test aims to check the accuracy of the metrology that can be obtained with a stereo sytem using the proposed technique for calibration. Figure 9 shows the experimental set-up used to measure the length of the black segment painted on the white board. The stereo algorithm matched the two extreme of the black segment by performing the epipolar transformation, described in section 7.2.

By simple triangulation, it is then possible to recover the absolute length of the black segment. The precision of the optical measure of the length of the black segment depends on its attitude in space. When the segment is almost orthogonal to the optical axis of the cameras we have a high precision which decreases when the segment tends to align parallel to the optical axes. In order to check the possibility of systematic error, we compared the absolute distance between each end point of the segment and one of the viewing cameras. Table 3 reproduces the obtained results. The results of these tests shown that our calibration procedure is suitable for a good stereo metrology [12], provided that the viewed objects are in focus and have an appropriate attitude in space.

8 Discussion

The proposed algorithm for camera calibration complements those already presented in the literature [4, 6, 12]. It has been tested and is extensively used in several national laboratories and provides calibration data with an accuracy adequate for metrology and stereoscopic vision. The idea of using vanishing points to locate objects in space or to recover the extrinsic parameters of a pair of images has already been suggested [14–16]. A modified procedure for measuring the extrinsic parameters, which takes advantage of the *a priori* information on the class of images to be processed in the edge-detection step, runs on a SUN 3 in

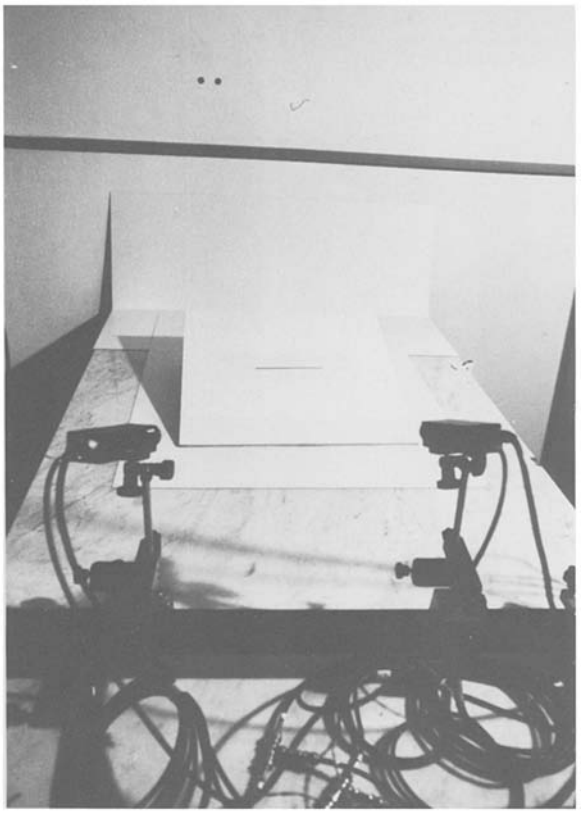

Fig. 9. The experimental set-up used to compute the length of a straight line segment. Two cameras are pointing towards a black segmer drawn on a white board.

Table 3. In each row are reported the length of the segment directly measured, the computed length, and the distance of an end point of the segment from the left camera measured directly and computed. The first five rows refers to segments almost orthogonal, and the last five rows to segments almost parallel, to the optical axes of the cameras.

	Measured Length	Computed Length	Measured Distance	Computed Distance
Almost orthogonal	123 ± 1 mm	122.5 ± 0.5 mm	425 ± 1 mm	418 ± 1.5 mm
	132 ± 1 mm	131 ± 2 mm	981 ± 1 mm	975 ± 2.5 mm
	137 ± 1 mm	136.5 ± 0.5 mm	1152 ± 1 mm	1158 ± 2 mm
	238 ± 1 mm	237 ± 0.1 mm	373 ± 1 mm	370 ± 3 mm
	245 ± 1 mm	246.0 ± 1 mm	788 ± 1 mm	795 ± 1 mm
Almost parallel	220 ± 1 mm	231 ± 5 mm	512 ± 1 mm	520 ± 8 mm
	250 ± 1 mm	244 ± 3.8 mm	483 ± 1 mm	470 ± 10 mm
	286 ± 1 mm	275 ± 4 mm	790 ± 1 mm	801 ± 7 mm
	342 ± 1 mm	325 ± 6 mm	598 ± 1 mm	585 ± 7 mm
	375 ± 1 mm	388 ± 5.1 mm	1150 ± 1 mm	1190 ± 5.1 mm

a few seconds. Therefore, this technique, which does not require the solution of a nonlinear system of equations, can be performed almost in real time.

Finally, let us make some remarks on the proposed technique suggested by extensive experimentation.

Intrinsic Parameters

The four intrinsic parameters, f, k_1/k_2, a, and b can be measured with varying precision, and their influence on the accuracy of the stereo system is also different. The ratio k_1/k_2 is measured manually off-line with very high precision and the figure obtained is in very good agreement with the value obtained from the data sheets of the instrumentation. The value of the focal length was also measured with different procedures, not described here, always with consistent results. The estimate of the location of the intersection of the optical axis with the image plane was not critical. Indeed, by performing the full calibration procedure with erroneous values of a and b (about 10–20%) we were still able to obtain good results in tests 2 and 3 of section 7. The reason for this behavior is due to the sequential estimates of intrinsic and extrinsic parameters: errors in a and b are compensated by different values of the translation $\mathbf{T}$.

Extrinsic Parameters

In the proposed procedure the rotation matrix $\mathbf{R}$ is estimated before the translation and has the very simple expression of equation (10). Moreover, error analysis show that the proposed recovery of $\mathbf{R}$ is very robust against inaccuracy of the localization of vanishing points. This property is the major reason for the practicality of the proposed technique. The estimation of the translation $\mathbf{T}$ is then obtained by a simple triangulation.

The proposed technique is definitely adequate for stereo vision and optical metrology. This is, in essence, the ultimate test for the proposed technique. The main practical advantage over previous aproaches [4, 13] is its simplicity.

Acknowledgment

We are indebted to A. Verri, F. Girosi, and T. Poggio for very helpful discussions. E. DeMicheli read several versions of the manuscript and made useful suggestions. Cristina Rosati kindly typed the manuscript and Clive Prestt checked the English. This work was supported by the ESPRIT project P940.

References

1. R.D. Arnold and T.O. Binford, "Geometric constraints in stereo vision," *Proc. SPIE*, San Diego, 1980.
2. H.H. Baker and T.O. Binford, "Depth from edge- and intensity-based stereo," *Proc. 7th Intern. Joint Conf. Artif. Intell.*, Vancouver, pp. 631–636, 1981.
3. R.C. Bolles, H.H. Baker, and D.H. Marimont, "Epipolar-plane image analysis: An approach to determining structure from motion," *Intern. Comput. Vision* 1:7–55, 1986.
4. O.D. Faugeras and G. Toscani, "Camera calibration for 3D computer vision," *Proc. Intern. Workshop Mach. Vision Mach. Intell.*, Tokyo, 1987.
5. O.D. Faugeras and G. Toscani, "The calibration problem for stereo," *Proc. IEEE Conf. Comput. Vision Pattern Recog.*, Miami, 1987.
6. R. Lenz and R.Y. Tsai, "Techniques for calibration of the scale factor and image center for high accuracy 3D machine vision metrology," *Proc. IEEE Intern. Conf. Robotics Automation*, Raleigh, NC, 1987.
7. M.J. Magee and J.K. Aggarwal, "Determining vanishing points from perspective images," *Comput. Vision, Graph., Image Process.*, 26:256-267, 1984.
8. Y. Ohta and T. Kanade, "Stereo by intra- and inter-scanline search using dynamic programming," Carnegie-Mellon University Tech. Rep. CMU-CS, 1983.
9. L. Pacioli, "De Divina Proportione". Reprint of *Codice Ambrosiano*, Fontes Ambrosiani, 31, Milano, Italy, 1956.
10. Piero della Francesca, *De Prospectiva Pingendi*, Edizione Critica, ed. Sansoni, Firenze, Italy, 1942.
11. A.N. Tikhonov and V.Y. Arsenin, *Solutions of Ill-Posed Problems*, Halsted Press, New York, 1977.
12. R.Y. Tsai, "A versatile camera calibration technique for high accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," IBM Res. Rpt. RC11413, 1985.
13. R.Y. Tsai and R. Lenz, "A new technique for fully autonomous and efficient 3D robotics hand/eye calibration," *Proc. 4th Intern. Symp. Robotics Res.*, Santa Cruz, 1987.
14. B. Odone, B. Caprile, F. Girosi, and V. Torre, "On the understanding of motion of rigid objects," no. 5, Intelligent Robots and Computer Vision; *Proc. SPIE 726*, 1987.
15. G. Nagy and A. Falsafi, "Using vanishing points to locate objects with six degrees of freedom," *Pattern Recognition in Practice III*, Amsterdam, May 1988.
16. B. Jiang, "Camera calibration technique using vanishing point concept," IEEE CH 2503-1/87/000-0640, 1987.