Building, Registrating, and Fusing Noisy Visual Maps

Nicholas Ayache
Olivier D. Faugeras

INRIA
Domaine de Voluceau, Rocquencourt
BP 105, 78153 Le Chesnay, France

Abstract

This paper deals with the problem of building three-dimensional descriptions (we call them visual maps*) of the environment of a mobile robot using passive vision. These maps are local (i.e., attached to specific frames of reference). Since noise is present, they incorporate information about the geometry of the environment and about the uncertainty of the parameters defining the geometry. This geometric uncertainty is directly related to its source (i.e., sensor uncertainty). We show how visual maps corresponding to different positions of the robot can be registered to compute a better estimate of its displacement between the various viewpoint positions, assuming an otherwise static environment. We use these estimates to fuse the different visual maps and reduce locally the uncertainty of the geometric primitives which have found correspondents in other maps. We propose to perform these three tasks (building, registrating, and fusing visual maps) within the general framework of extended Kalman filtering, which allows efficient combination of measurements in the presence of noise.*

1. Introduction

The problem of dealing with noise in three-dimensional vision and mobile robots is one of the first to be tackled in order to make both things useful. We believe that it cannot be engineered away and that its solution has to be found, first, by representing explicitly the uncertainty in the world model used by the robot and, second, by combining a large number of measurements and/or sensors.

In this article, we propose a partial solution along those lines in the case where passive stereo is used to collect three-dimensional information. This work continues the one presented in Ayache et al. (1985) and Faugeras et al. (1986) and is connected to that of Brooks (1985), Bolle and Cooper (1985), Laumond and Chatila (1985), Crowley (1986), Smith and Cheeseman (1986), and Durrant-Whyte (1986).

The goals to be reached are the establishment of a three-dimensional description of the environment in which the vehicle moves, which includes both geometric information and information about the uncertainty attached to the corresponding geometric primitives. This description can then be used for various tasks, such as the definition of a sensing strategy—where to look next in order to increase the accuracy of the model—and a navigation strategy—how to go from here to there given our present state of knowledge; it can also be used to detect changes in the environment or recognize places. These applications are not incorporated in this paper.

The key points of our approach are the use of a powerful tool for dealing with large numbers of noisy measures, the extended Kalman filter (Darmon 1982; Jazwinsky 1970), which we found extremely useful in some of our previous work (Ayache and Faugeras 1986; Faugeras and Hebert 1986), and the way we represent three-dimensional rotations to relate frames of reference.

The latest developments of this work and new results are presented in Ayache and Faugeras (1988) and Ayache (1988).

2. Linearizing the Problem

In all the cases we discuss, we deal with an observation $\mathbf{x}$ in $\mathbf{R}^m$ that depends on a parameter $\mathbf{a}$ in $\mathbf{R}^n$ in a nonlinear fashion that can be expressed as a relation $\mathbf{f}(\mathbf{x}, \mathbf{a}) = \mathbf{0}$, where $\mathbf{f}$ maps $\mathbf{R}^m \times \mathbf{R}^n$ into $\mathbf{R}^p$. We assume

The International Journal of Robotics Research,
Vol. 7, No. 6, December 1988,

that the observation $\mathbf{x}$ is corrupted with noise, which we model as an additive zero mean Gaussian noise:

$$\mathbf{x} = \mathbf{x}' + \epsilon \quad \text{with } E(\epsilon) = \mathbf{0} \text{ and } E(\epsilon\epsilon^t) = \Lambda.$$

The problem is, given a number of observations $\mathbf{x}_i$, to find the parameter vector $\mathbf{a}$ that best satisfies the relations $\mathbf{f}_i(\mathbf{x}_i, \mathbf{a}) = \mathbf{0}$. "Best" is to be made more precise in a moment.

Let us drop the i indices for a while. Supposing that we know a "good" estimate $\mathbf{a}^*$ of $\mathbf{a}$, we can use the idea of linearization and expand $\mathbf{f}$ in the vicinity $(\mathbf{x}', \mathbf{a})$:

$$\mathbf{f}(\mathbf{x}', \mathbf{a}) = \mathbf{0} \approx \mathbf{f}(\mathbf{x}, \mathbf{a}^*) + \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}, \mathbf{a}^*)(\mathbf{x}' - \mathbf{x}) + \frac{\partial \mathbf{f}}{\partial \mathbf{a}}(\mathbf{x}, \mathbf{a}^*)(\mathbf{a} - \mathbf{a}^*).$$

As usual, $\partial \mathbf{f}/\partial \mathbf{x}$ is a $p \times m$ matrix and $\partial \mathbf{f}/\partial \mathbf{a}$ is a $p \times n$ matrix. This expression can be rewritten as (dropping the $\approx$ sign):

$$-\mathbf{f}(\mathbf{x}, \mathbf{a}^*) + \frac{\partial \mathbf{f}}{\partial \mathbf{a}}(\mathbf{x}, \mathbf{a}^*)\,\mathbf{a}^* = \frac{\partial \mathbf{f}}{\partial \mathbf{a}}(\mathbf{x}, \mathbf{a}^*)\,\mathbf{a} - \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x} - \mathbf{a}^*)\,\epsilon,$$

which is a linear measurement equation:

$$\mathbf{y} = \mathbf{M}\mathbf{a} + \mathbf{u},$$

where

$$\mathbf{y} = -\mathbf{f}(\mathbf{x}, \mathbf{a}^*) + \frac{\partial \mathbf{f}}{\partial \mathbf{a}}(\mathbf{x}, \mathbf{a}^*)\,\mathbf{a}^*,$$

$$M = \frac{\partial \mathbf{f}}{\partial \mathbf{a}}(\mathbf{x}, \mathbf{a}^*) \quad \text{and} \quad \mathbf{u} = -\frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}, \mathbf{a}^*)\,\epsilon.$$

Notice that $\mathbf{y}$ and $\mathbf{M}$ are known because we know $\mathbf{f}$, $\mathbf{x}$, and $\mathbf{a}^*$; since we also know Λ, we know $\mathbf{u}$'s second-order statistics:

$$E(\mathbf{u}) = \mathbf{0},$$

$$W = E(\mathbf{u}\mathbf{u}^t) = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}, \mathbf{a}^*)\,\Lambda\,\frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}, \mathbf{a}^*)^t.$$

Let us now come back to our $\mathbf{n}$ observations $\mathbf{x}_i$; for each of them, knowing the function $\mathbf{f}_i$ and the covariance Λ_i of $\mathbf{x}_i$, one can form a linear measurement equation

$$\mathbf{y}_i = \mathbf{M}_i\mathbf{a} + \mathbf{u}_i$$

and compute the noise covariance matrix $\mathbf{W}_i$ attached to it. If we start with an initial estimate $\hat{\mathbf{a}}_0$ of $\mathbf{a}$ and its associated covariance matrix $\mathbf{S}_0 = E((\hat{\mathbf{a}}_0 - \mathbf{a})(\hat{\mathbf{a}}_0 - \mathbf{a})^t)$, we can use the Kalman filtering approach to deduce recursively an estimate $\hat{\mathbf{a}}_n$ of $\mathbf{a}$ and its covariance matrix $\mathbf{S}_n((\hat{\mathbf{a}}_n - \mathbf{a})(\hat{\mathbf{a}}_n - \mathbf{a})^t)$ after taking into account n observations. The corresponding recursive equations are the standard Kalman equations and are given in Appendix A. We can now give a precise meaning to the word "best." $\hat{\mathbf{a}}_n$ is the parameter vector that minimizes the criterion

$$(\mathbf{a} - \hat{\mathbf{a}}_0)^t\mathbf{S}_0^{-1}(\mathbf{a} - \hat{\mathbf{a}}_0) + \sum_{i=1}^{n}(\mathbf{y}_i - \mathbf{M}_i\mathbf{a})^t\,\mathbf{W}_i^{-1}(\mathbf{y}_i - \mathbf{M}_i\mathbf{a}).$$

This equation is important, because it shows how the Kalman filtering explicitly takes into account noise in the measurements and weighs them accordingly. The more noise we have on the ith measurement, the "smaller" the inverse covariance matrix $\mathbf{W}_i^{-1}$ is, and therefore the less the ith term in the above criterion contributes to the final estimate.

Let us now study in detail the application of this technique to some previously mentioned problems.

3. Stereo Reconstruction

In a standard, two-camera, stereo system, such as the one depicted in Fig. 1, we can relate the (x, y, z) coordinate space to the (u_1, v_1) and (u_2, v_2) retina spaces linearly in projective spaces:

$$\begin{bmatrix} s_i & u_i \\ s_i & v_i \\ s_i & \end{bmatrix} = \mathbf{P}_i \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}, \qquad i = 1, 2,$$

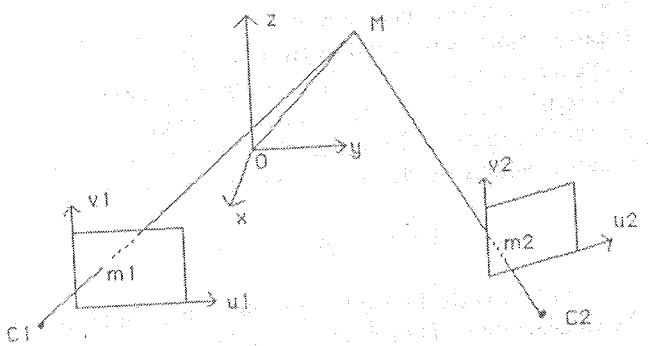


Fig. 1. Geometry of a standard, two-camera, stereo system.

where matrices $\mathbf{P}_i$ are obtained by calibration (Faugeras and Toscani 1986). When we match a point of coordinates (u_1, v_1) in retina 1 to a point of coordinates (u_2, v_2) in retina 2, we can compute by triangulation the coordinates (x, y, z) of the corresponding point in three-dimensional space. To cast this in the previous formalism, let us write (dropping the indexes for the moment)

$$\mathbf{x} = [u, v]^t, \qquad \mathbf{a} = [x, y, z]^t = \mathbf{OM}.$$

Rewriting the 3×4 matrix $\mathbf{P}$ as

$$\mathbf{P} = \begin{bmatrix} \mathbf{l}_1 & l_{14} \\ \mathbf{l}_2 & l_{24} \\ \mathbf{l}_3 & l_{34} \end{bmatrix},$$

where $\mathbf{l}_1$, $\mathbf{l}_2$, $\mathbf{l}_3$ are 3×1 row vectors, and eliminating s, we have the following relationships:

$$(\mathbf{l}_3\mathbf{OM} + l_{34})u - \mathbf{l}_1\mathbf{OM} - l_{14} = 0,$$
$$(\mathbf{l}_3\mathbf{OM} + l_{34})v - \mathbf{l}_2\mathbf{OM} - l_{24} = 0,$$

Therefore,

$$\mathbf{f}(\mathbf{x}, \mathbf{a}) = \begin{bmatrix} (\mathbf{l}_3\mathbf{OM} + l_{34})\, u - \mathbf{l}_1\mathbf{OM} - l_{14} \\ (\mathbf{l}_3\mathbf{OM} + l_{34})\, v - \mathbf{l}_2\mathbf{OM} - l_{24} \end{bmatrix}. \qquad (2)$$

From the above formalism we can deduce an estimate $\hat{\mathbf{a}}_2$ of $\mathbf{a}$ after two measurements (i.e., of the position in three-dimensional space of the point M) as well as the covariance matrix of this estimate, which is an indication of the uncertainty of our knowledge. This is a function of our uncertainty on the pixel coordinates (u_1, v_1) and (u_2, v_2) and of our initial estimate $\hat{\mathbf{a}}_0$ of the position of M and its covariance matrix $\mathbf{S}_0$. The equations are in Appendix B. In practice, $\mathbf{S}_0$ is diagonal with very large diagonal terms, so the first term in Eq. (1) is very small.

This shows how the uncertainty on points obtained by stereo can be computed. In practice, we consider points, lines, and planes. The uncertainty on these primitives depends first on how they have been extracted and, second, on how they are represented. Let us look at representations first.

A possible way to represent lines is to use a vector parallel to the line and a point on the line. A line D is represented by two vectors $(\mathbf{l}, \mathbf{m})$ defined as follows. Given two points M_1 and M_2 on L, we have

$$\mathbf{l} = \mathbf{OM}_2 - \mathbf{OM}_1, \qquad \mathbf{m} = (\mathbf{OM}_1 + \mathbf{OM}_2)/2.$$

Planes are represented by their normal $\mathbf{n}$ and their distance to the origin d. See Faugeras and Hebert (1986) for an application of this representation to the recognition and localization of three-dimensional objects.

The description built by the stereo system consists of two parts. It produces, first, the parameters of the geometric elements represented and second, the uncertainty on these parameters. This uncertainty is represented by a set of covariance matrices.

We have treated the case of points in this section. The case of lines can be deduced very simply from the above representation, as shown in the next section. The interested reader can find an analysis for the case of planes in Faugeras and Lustman (1986).

The resulting description of the environment, including the geometric and uncertainty aspects is called a *realistic uncertain description of the environment.*

4. Registrating Stereo Pairs

The problem is the following. Suppose that the analysis of a first stereo pair yields a partial three-dimensional description of the environment in terms of points,

lines, and planes, each with some uncertainty attached to it. The problem of obtaining this uncertainty has been treated in the previous section for points.

Now suppose that the vehicle on which the cameras are mounted moves (let us say in three-dimensions) by an imperfectly known displacement $\mathbf{D}$. We then acquire a second stereo pair, analyze it, and obtain another three-dimensional description of another part of the environment. If the displacement $\mathbf{D}$ is not too large, it is likely that some of the geometric primitives identified in position 1 will also be identified in position 2. By matching such primitives we should be able to recover a better estimate of the displacement $\mathbf{D}$ and (this is done in Section 6) to construct a better estimate of the position and orientation of these primitives which have been identified in the two positions; i.e., we should be able to improve the description of the environment.

We must also say something about the way we represent three-dimensional displacements. Every such displacement can be decomposed in an infinite number of ways as the product of a translation characterized by a vector $\mathbf{t}$ and a rotation $\mathbf{R}$ characterized by its angle θ, its origin, and its axis $\mathbf{u}$ (a unit vector). Fixing the origin of the rotation makes the decomposition unique.

For our next purpose we can therefore consider the group of three-dimensional displacements as the product of the group SO_3 of rotations and the group of translations $\mathbf{R}^3$. How we parameterize SO_3 is also important. For an excellent review of the different parameterizations of SO_3, see Stuelpnagel (1964). In our previous work (Faugeras and Hebert 1986) we used quaternions. This has the disadvantage of imposing the constraint that the quaternions dealt with are of unit norm. In this article we propose to use the exponential form of a rotation matrix $\mathbf{R}$. Indeed, for every orthogonal matrix $\mathbf{R}$, there exists a unique antisymmetric matrix $\mathbf{H}$ such that $\mathbf{R} = e^{\mathbf{H}}$, where matrix exponentials are defined as usual as $e^{\mathbf{H}} = \mathbf{I} + \mathbf{H}/1! + \mathbf{H}^2/2! + \cdots$ The matrix $\mathbf{H}$ can be written as

$$\mathbf{H} = \begin{bmatrix} 0 & -c & b \\ c & 0 & -a \\ -b & a & 0 \end{bmatrix}.$$

The three-dimensional vector $\mathbf{r} = [a, b, c]^t$ has some useful properties. Its direction is that of the axis of rotation, and its squared norm is equal to the rotation angle squared. Proofs of that and other properties of this representation are presented in Appendix C. Moreover, matrix $\mathbf{H}$ represents the cross product with vector $\mathbf{r}$, which we denote by $c(\mathbf{r})$. By this we mean

$$\mathbf{H}\mathbf{x} = c(\mathbf{r})\mathbf{x} = \mathbf{r} \wedge \mathbf{x}.$$

We can now deal with the problem of estimating the displacement $\mathbf{D}$ and its uncertainty by matching geometric primitives detected in positions 1 and 2 of the vehicle. Let us start with points first. If the same physical point is represented in the coordinate system associated with the first position by $\mathbf{OM}$ and in the coordinate system associated with the second position by $\mathbf{O'M'}$ then

$$\mathbf{O'M'} - \mathbf{ROM} - \mathbf{t} = \mathbf{0}. \tag{3}$$

This equation is of the form $\mathbf{f}(\mathbf{x}, \mathbf{a})$, if we let $\mathbf{x} = [\mathbf{O'M'}^t, \mathbf{OM}^t]^t$ and $\mathbf{a} = [\mathbf{r}^t, \mathbf{t}^t]^t$ with $\mathbf{r}$ defined as above. Therefore, the previous formalism can again be used and an estimate of both $\mathbf{r}$ and $\mathbf{t}$ can be built by matching a number of points in positions 1 and 2. The uncertainty on $\mathbf{r}$ and $\mathbf{t}$ can also be computed if we have a model of the uncertainty on $\mathbf{OM}$ and $\mathbf{O'M'}$, which we have after Section 3. The equations are presented in Appendix D.

The problem of matching straight lines is very similar. Let the two lines be defined by two points (M_1M_2) and $(M'_1M'_2)$ defining representations $(\mathbf{l}, \mathbf{m})$ and $(\mathbf{l}', \mathbf{m}')$. When a line is submitted to a rotation $\mathbf{R}$ and a translation $\mathbf{t}$, it is fairly easy to show that its representation becomes $(\mathbf{Rl}, \mathbf{Rm} + \mathbf{t})$. We can then write that

$$\mathbf{l}' \times \mathbf{Rl} = \mathbf{0} \tag{4}$$

and

$$\mathbf{Rl} \times (\mathbf{m}' - \mathbf{Rm} - \mathbf{t}) = \mathbf{0}$$

or, using (3),

$$\mathbf{l}' \times (\mathbf{m}' - \mathbf{Rm} - \mathbf{t}) = \mathbf{0}. \tag{5}$$

The geometric interpretation of Eq. (3) is that the

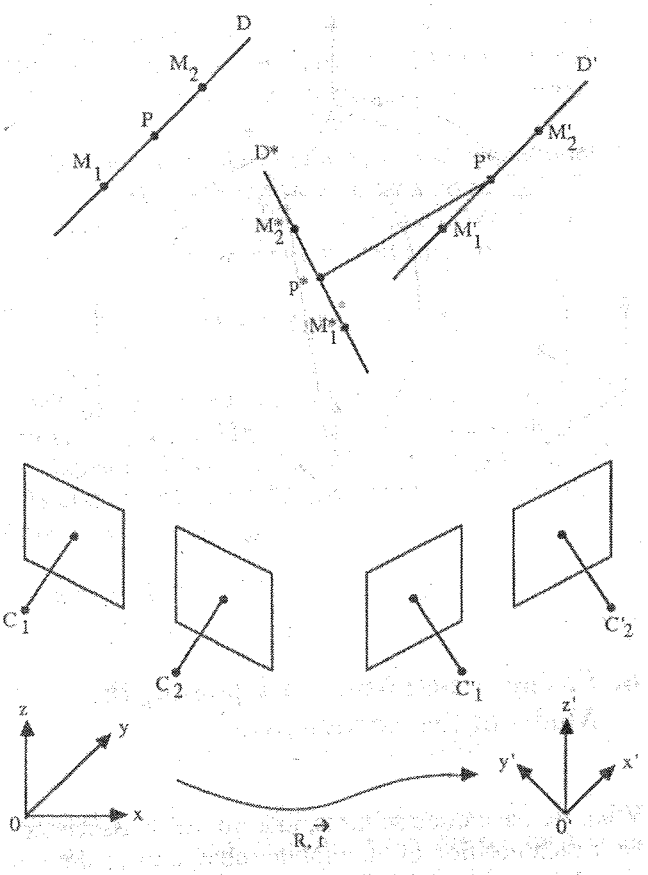


Fig. 2. D is a line reconstructed by stereo in the coordinate system (0xyz). D′ is a line reconstructed by stereo in the coordinate system (0′x′y′z′) and D is D displaced by the rotation R and the translation t. We can write that D* is the same as D′.*

transformed line is parallel to the second line. The geometric interpretation of Eq. (4) is that the line joining the midpoints of the transformed segment and the second segment is parallel to the second line (see Fig. 2).

Letting $\mathbf{x} = (\mathbf{l}'^{\mathrm{t}}, \mathbf{m}'^{\mathrm{t}}, \mathbf{l}^{\mathrm{t}}, \mathbf{m}^{\mathrm{t}})$ and $\mathbf{a}$ as before, this is again of the form $\mathbf{f}(\mathbf{x}, \mathbf{a}) = \mathbf{0}$, where $\mathbf{f}(\mathbf{x}, \mathbf{a})$ is a vector. But in fact the measurement equations produced by a cross product are not independent: the knowledge of two of them is sufficient to derive the third one. Therefore one must keep only four measurements out of six, the first two coming from the first cross product, the last two coming from the second cross product. Then $\mathbf{f}(\mathbf{x}, \mathbf{a})$ becomes a 4-vector. The equations are presented in Appendix E.

The case of planes can be treated similarly. Letting the same physical plane be represented by $(\mathbf{n}', d')$ and $(\mathbf{n}, d)$ in coordinate systems 1 and 2, we can write

$$\mathbf{n}' - \mathbf{R}\mathbf{n} = \mathbf{0} \tag{6}$$

and

$$d' - d + \mathbf{t}^{\mathrm{t}}\mathbf{R}\mathbf{n} = 0$$

or

$$d' - d - \mathbf{t}^{\mathrm{t}}\mathbf{n}' = 0. \tag{7}$$

Letting $x = (\mathbf{n}'^{\mathrm{t}}, \mathbf{n}^{\mathrm{t}}, d', d)^{\mathrm{t}}$ and $\mathbf{a}$ as before, we can again put this equation in the form $\mathbf{f}(\mathbf{x}, \mathbf{a}) = \mathbf{0}$. Equations are presented in Appendix F.

5. Registrating Stereo Pairs and Images

In order to avoid solving the stereo matching problem too often, it may be useful, once a three-dimensional estimation of the scene has been constructed, to track the projections in one image of the geometric features used to compute the three-dimensional displacement. This is similar in spirit to the work of Lowe, who does not use our formalism (Lowe 1985). Let us first derive the case of points. We assume that we know the image coordinates u', v' of point M' such that $\mathbf{O'M'} = \mathbf{ROM} + \mathbf{t}$. Therefore, letting $\mathbf{x} = [\mathbf{OM}^{\mathrm{t}}, u', v']^{\mathrm{t}}$ and $\mathbf{a} = [\mathbf{r}^{\mathrm{t}}, \mathbf{t}^{\mathrm{t}}]^{\mathrm{t}}$, we can use a combination of Eqs. (2) and (3):

$$\mathbf{f}(\mathbf{x}, \mathbf{a}) = \begin{cases} (\mathbf{I}_3(\mathbf{ROM} + \mathbf{t}) + l_{34})u' - \mathbf{l}_1(\mathbf{ROM} + \mathbf{t}) - l_{14}, \\ (\mathbf{I}_3(\mathbf{ROM} + \mathbf{t}) + l_{34})v' - \mathbf{l}_2(\mathbf{ROM} + \mathbf{t}) - l_{24}. \end{cases} \tag{8}$$

The equations for $\partial\mathbf{f}/\partial\mathbf{x}$ and $\partial\mathbf{f}/\partial\mathbf{a}$ are presented in Appendix G.

Let us now derive the case of lines. Let D be a three-dimensional line defined by two points M_1 and M_2. After rotation and translation, the transformed line D' is defined by M'_1 and M'_2 such that $\mathbf{O'M}'_1 = \mathbf{ROM}_1 + \mathbf{t}$ and $\mathbf{O'M}'_2 = \mathbf{ROM}_2 + \mathbf{t}$. The projection d' of D' in the retina of one of the cameras is defined by the two points m'_1 and m'_2, projections of M'_1 and M'_2. If we denote by $[U'_1, V'_1, T'_1]$ and $[U'_2, V'_2, T'_2]$ the projective coordinates of the points m'_1 and m'_2, the projective equation of the line d' is given by the determinant

$$\begin{vmatrix} U & U'_1 & U'_2 \\ V & V'_1 & V'_2 \\ T & T'_1 & T'_2 \end{vmatrix}$$

Fig. 3. After the rotation R and translation $\vec{t}$, the line D in coordinate system (0xyz) becomes line D′ in coordinate system (0′x′y′z′). D′ projects as d′ in retina 1 and is matched to d″.

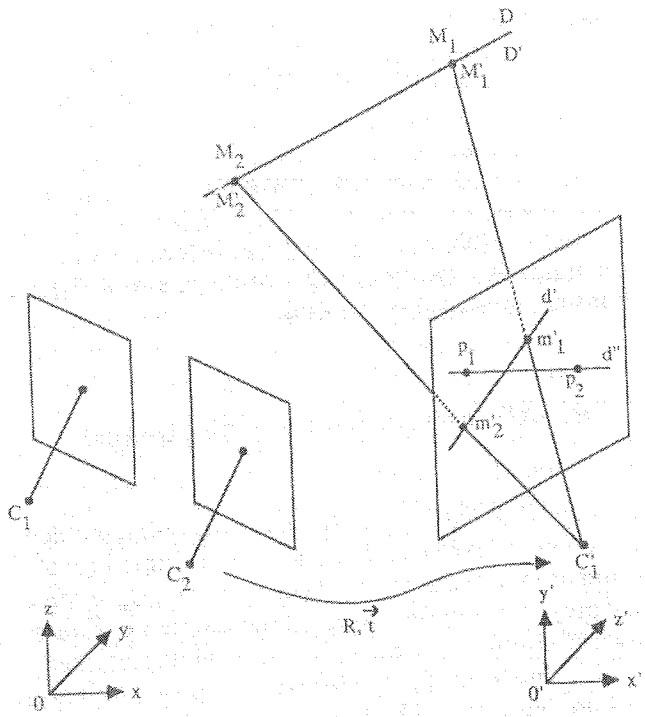
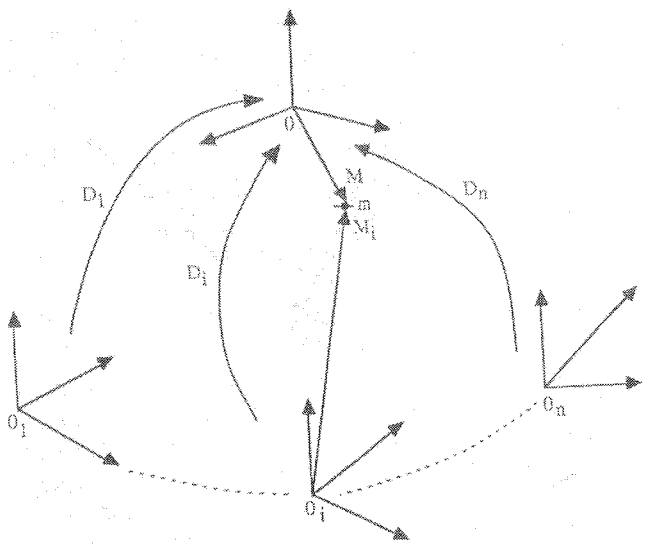


Fig. 4. The physical point m is represented by M, $M_1, \ldots, M_n$ in the various coordinate systems.



or

$$U(V'_1T'_2 - V'_2T'_1) + V(T'_1U'_2 - T'_2U'_1) + T(U'_1V'_2 - U'_2V'_1) = 0.$$

If we match d' to a line d'' in the focal plane defined by two points p_1 and p_2 (see Fig. 3) with projective coordinates $[u_1, v_1, 1]$ and $[u_2, v_2, 1]$, the necessary and sufficient condition for the two lines to be the same is $\mathbf{h}_1 \times \mathbf{h}_2 = \mathbf{0}$, where $\mathbf{h}_1 = [V'_1T'_2 - V'_2T'_1, T'_1U'_2 - T'_2U'_1, U'_1V'_2 - U'_2V'_1]^t$ and $\mathbf{h}_2 = [v_1 - v_2, u_2 - u_1, u_1v_2 - u_2v_1]^t$. Now, with $\mathbf{x} = [\mathbf{OM}'^t_1, \mathbf{OM}'^t_2, u_1, v_1, u_2, v_2]^t$ and $\mathbf{a}$ as before, the measurement equation is

$$\mathbf{f}(\mathbf{x}, \mathbf{a}) = \mathbf{h}_1 \times \mathbf{h}_2 = \mathbf{0},$$

where the components of $\mathbf{h}_1$ can be easily computed as functions of $\mathbf{x}$ and $\mathbf{a}$. Here again, $\mathbf{f}(\mathbf{x}, \mathbf{a})$ is a 3-vector, but the measurements are correlated and only two measurements can be kept. The details are presented in Appendix G.

6. Fusing Visual Maps or Updating the Model of the Environment

What we have done so far is, first, to build, associated with each position of the mobile robot, a three-dimensional description of the environment in terms of its geometry (the positions and orientations of points, lines, and planes) and the uncertainty of the parameters describing these primitives. Each such description is attached to a local coordinate frame. Second, when there exist physical primitives which are common to two frames of reference, and we do not know exactly (perhaps even not at all) the relative position and orientation of the two frames, we have shown that, by matching primitives across frames, we were capable of building estimates of the three-dimensional transformation between the frames and a measure of the uncertainty of this transformation.

The last step is to close the loop and use this information to update, in each local frame, the description of the geometry and uncertainty of the primitives corresponding to parts of physical objects visible in another frame. This situation corresponds to that depicted in Fig. 4, where m is a physical point represented by the vector $\mathbf{O}_i\mathbf{M}_i$ and the covariance matrix

cov $\mathbf{M}_i$ in frame $\mathbf{F}_i$ ($i = 1, \ldots, n$), and by the vector $\mathbf{OM}$ and the covariance matrix cov $\mathbf{M}$ in a frame $\mathbf{F}$ in which we want to update the position and uncertainty of m.

Frame $\mathbf{F}$ is related to frame i by rigid displacements $\mathbf{D}_1, \ldots, \mathbf{D}_n$, each represented by a vector $[\mathbf{r}_i^t, \mathbf{t}_i^t]^t$ and its covariance matrix Λ_i in frame i. We therefore have the following n measurement equations:

$$\mathbf{OM} - \mathbf{R}_i\mathbf{O}_i\mathbf{M}_i - \mathbf{t}_i = \mathbf{0},$$

which are of the form $\mathbf{f}_i(\mathbf{x}, \mathbf{a}) = \mathbf{0}$ with $\mathbf{a} = [\mathbf{O}_n\mathbf{M}_n]$ and $\mathbf{x}_i = [\mathbf{r}_i^t, \mathbf{t}_i^t, \mathbf{O}_i\mathbf{M}_i^t]^t$. We can apply the artillery developed in Section 2 and obtain a new estimate of $\mathbf{OM}$ and cov $\mathbf{M}$. Therefore, using the notations of Appendix C, we get

$$\partial\mathbf{f}_i/\partial\mathbf{x} = [-\mathbf{K}(\mathbf{R}_i, \mathbf{O}_i\mathbf{M}_i) \quad -\mathbf{I} \quad -\mathbf{R}_i] \qquad \text{a } 3 \times 9 \text{ matrix}$$

and

$$\partial\mathbf{f}_i/\partial\mathbf{a} = \mathbf{I} \qquad \text{a } 3 \times 3 \text{ matrix.}$$

Similarly, in the case of lines, we could try to use Eqs. (4) and (5) to update the estimate of the representation of a given line and its associated uncertainty:

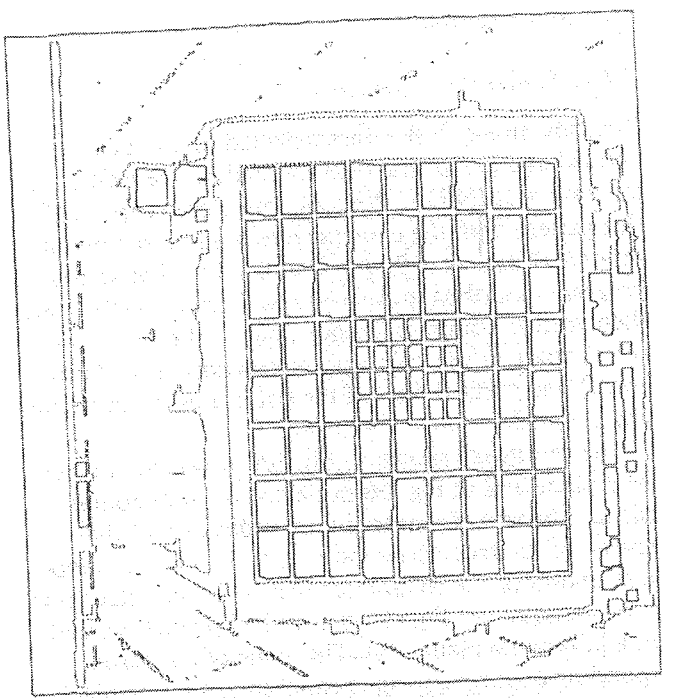
$$\mathbf{l} \times \mathbf{R}_i\mathbf{l}_i = \mathbf{0} \qquad \text{and} \qquad \mathbf{l} \times (\mathbf{m} - \mathbf{R}_i\mathbf{m}_i - \mathbf{t}_i) = \mathbf{0},$$

which are of the form $\mathbf{f}_i(\mathbf{x}, \mathbf{a}) = \mathbf{0}$ (with $\mathbf{a} = [\mathbf{l}^t, \mathbf{m}^t]^t$ and $\mathbf{x} = [\mathbf{r}_i^t, \mathbf{t}_i^t, \mathbf{l}_i^t, \mathbf{m}_i^t]^t$), keeping only four independent measurements out of six. But it is better to use minimal representations of primitives updated by the Kalman filter (i.e., four-parameter representations of lines). This is discussed and experimented with in Ayache and Faugeras (1987,1988).

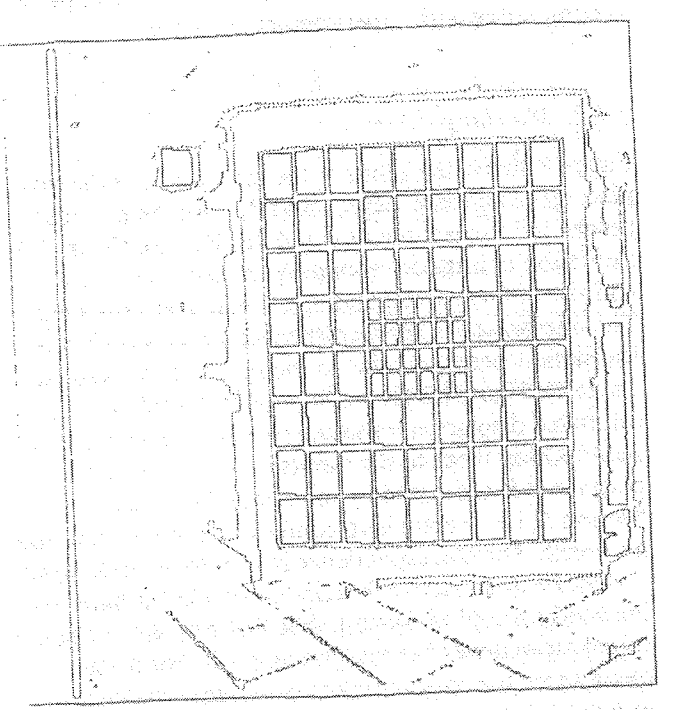
7. Implementation and Results

We have implemented the theoretical results developed in the previous sections and run the programs on a number of real sequences of stereo pairs.

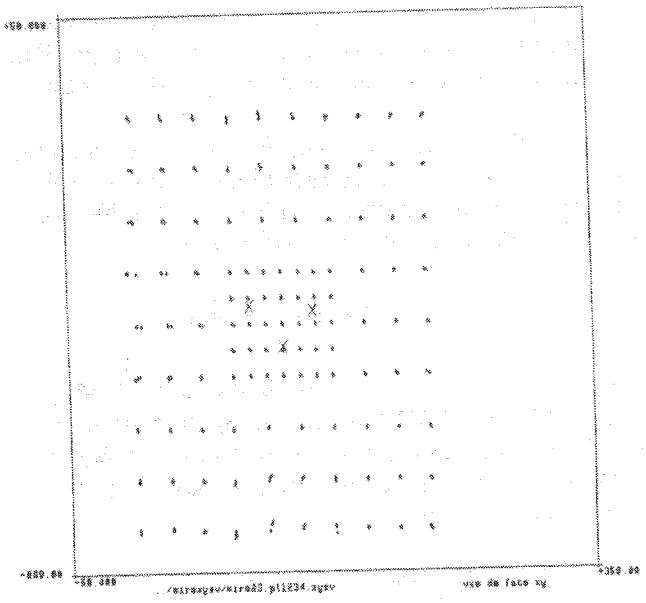
Fig. 5. Edges extracted from a stereo view of the grid pattern.



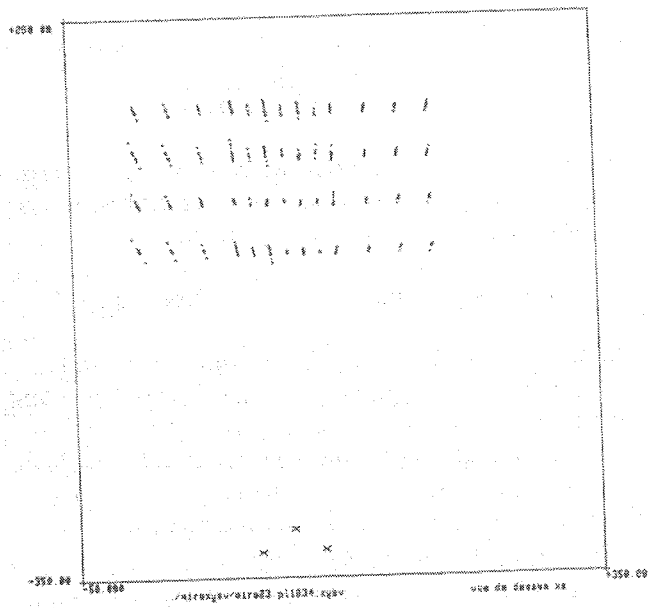
(a)



(b)



(a)



(b)

Fig. 6. Horizontal and vertical projection of the reconstructed intersections of the grid pattern observed in four positions.

7.1. Grid Example

7.1.1. Stereo Reconstruction

We show in Fig. 5 the lines detected in an image of vertical and horizontal lines painted on the wall of the laboratory at INRIA. We took four stereo triplets of this pattern with the cameras mounted on the mobile robot at distances of 3, 3.5, 4, and 4.5 m. The stereo program described in Ayache and Faverjon (1987) was then used to compute the positions in three-dimensional space of the points of intersection of the horizontal and vertical lines of the pattern using only cameras 2 and 3.

The results are shown in Fig. 6A projected in a vertical plane and in Fig. 6B projected in a horizontal plane. The sets of points give an indication of the spread of the reconstruction results. Figure 7 is a representation of the covariance matrices computed from the measurement Eq. (1) in a vertical plane (7a), and in a horizontal plane (7b). The value of the corresponding quadratic form is equal to 1, and the pixel noise was taken to be of variance 1 pixel. There is an excellent qualitative agreement between Fig. 6 and 7, indicating that the extended Kalman filtering approach is compatible with experimental evidence.

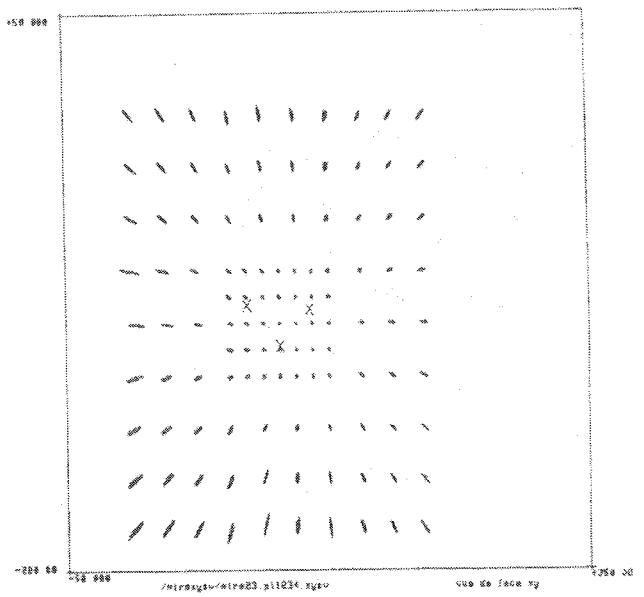
7.1.2. Matching

Figure 8 shows the result of the estimation of the displacement between views 1 and 2 using the previous points and measurement equation, Eq. (3). Figure 8A shows the points corresponding to view 1 as crosses and the points corresponding to view 2 as plus signs. The displacement between the two views is mostly a translation perpendicular to the plane of the wall (see Table 1). Figure 8B shows the result of applying the estimated displacement to the points in view 2 and superimposing them to the points in view 1. The displacement is estimated by the method described in Section 5; i.e., we are here combining a stereo pair and an image. The correspondence is seen to be quite good.

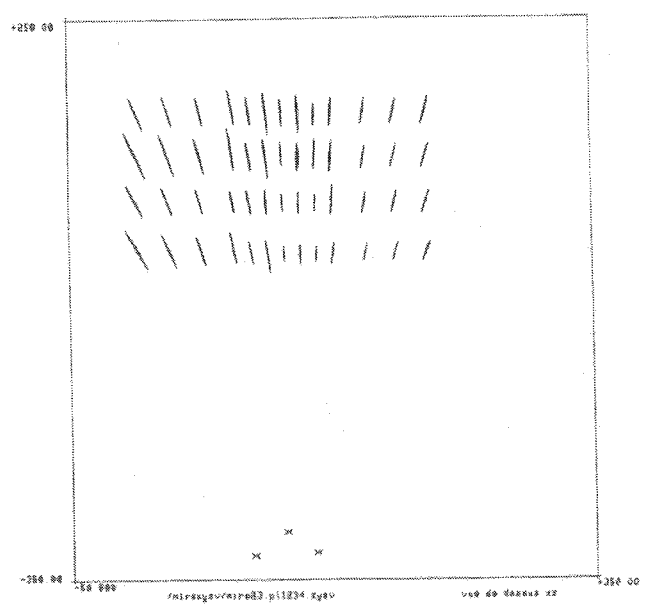
A more quantitative description of what is happening can be found in Table 1. The first row shows the initial estimates of the rotation and translation, the second row the corresponding covariance matrices, and the third and fourth rows the estimates found by running the extended Kalman filter twice over a set of 113 points. The fifth row shows the actual displacement between the first and fourth positions. The final covariance matrices are very small and are not shown.

Figure 9 shows a similar example where a rotation has been added to the translation. Figure 9A shows

Fig. 7. Covariance matrices attached to the points shown in Fig. 6.

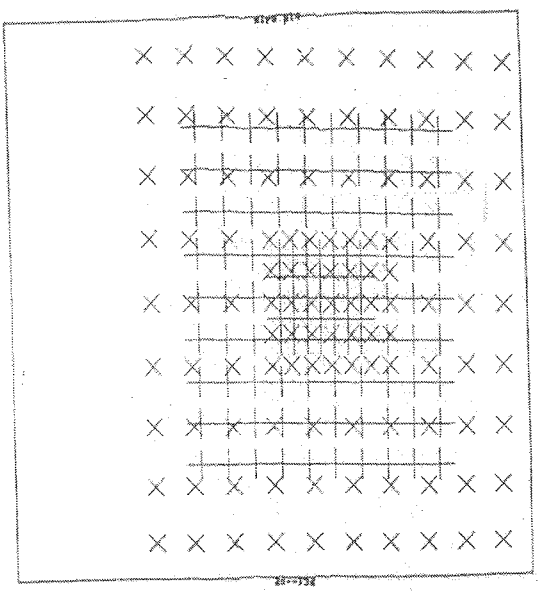


(a)

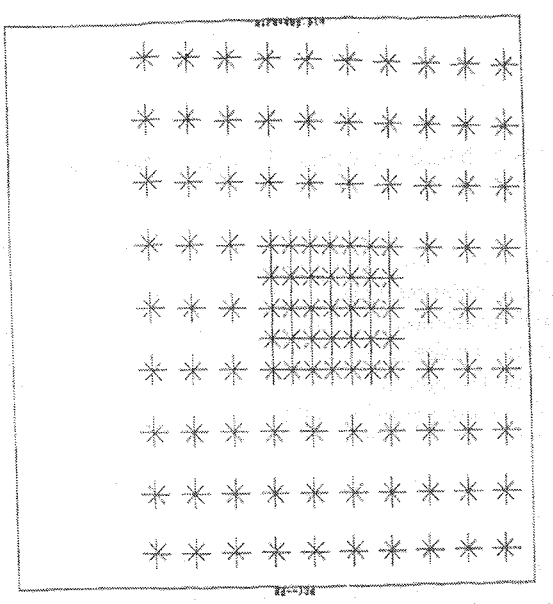


(b)

Fig. 8. Images of the grid pattern observed in two different positions before and after application of the estimated displacement.



(a)



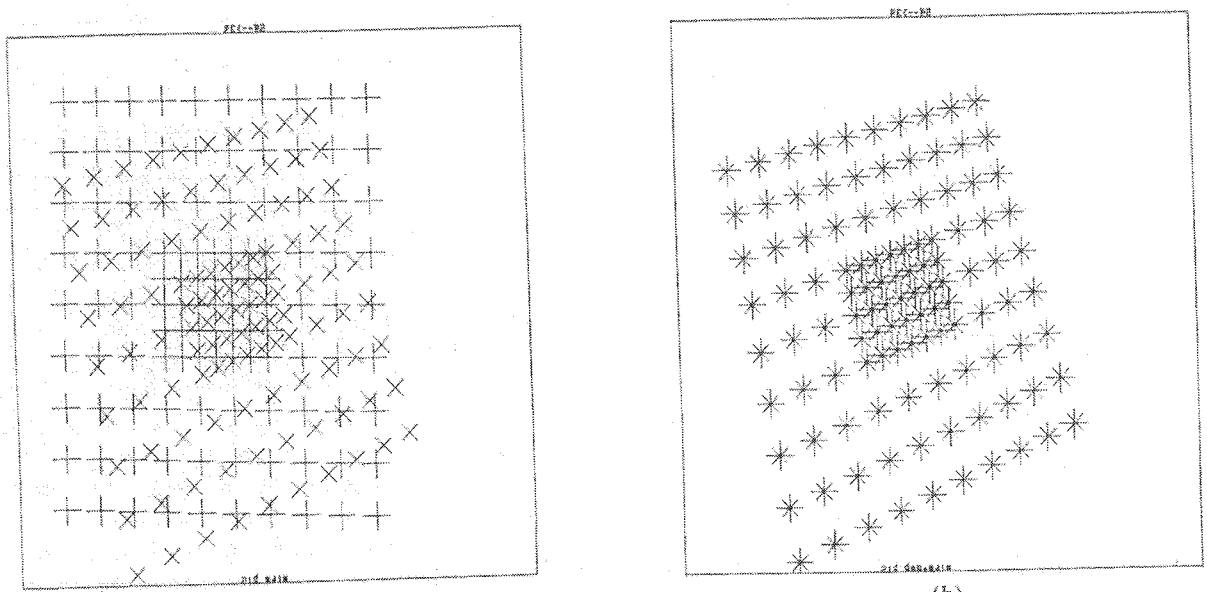
(b)

the points in views 1 and 2, and Fig. 9B shows the points after the estimated displacement has been applied to the points in view 2. In both cases, similar results for the estimated displacement are obtained when stereo pairs are combined. Table 2 details the results in the same format as Table 1.

7.1.3. Fusion

Figures 10 and 11 show the results of the integration of two different kinds of information. First, the points are reconstructed in each plane using the third camera (this is just adding one more measurement equation of

Fig. 9. Images of the grid pattern observed in two different positions before and after application of another estimated displacement.



(a) (b)

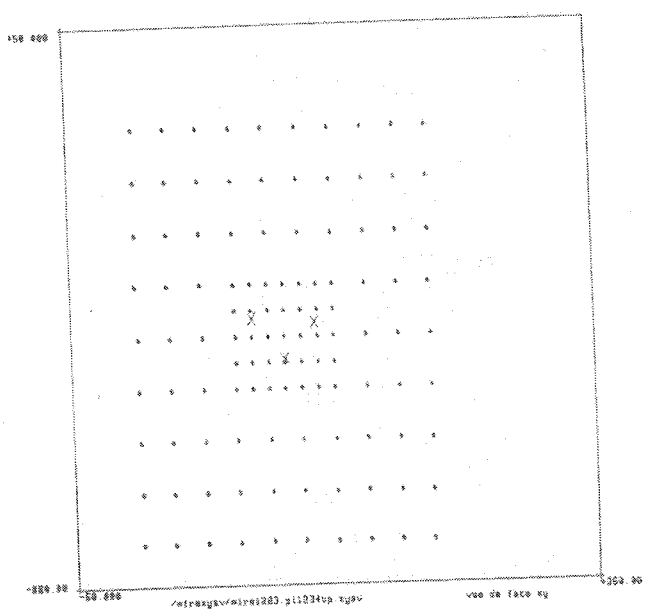
Table 1. Motion between Two Grid Patterns

	r_x (rad)	r_y (rad)	r_z (rad)	t_x (cm)	t_y (cm)	t_z (cm)	Angle (°)
Motion initial estimate	0	0	0	0	0	0	0
Initial covariance	1.0	1.0	1.0	100	100	100	
Motion iteration 1	0.009	−0.002	−0.000	4.41	1.94	−154.01	1.40
Motion iteration 2	0.002	−0.001	−0.002	0.88	1.03	−150.39	0.21
Real motion manually computed	0	0	0	0	0	−150.00	0

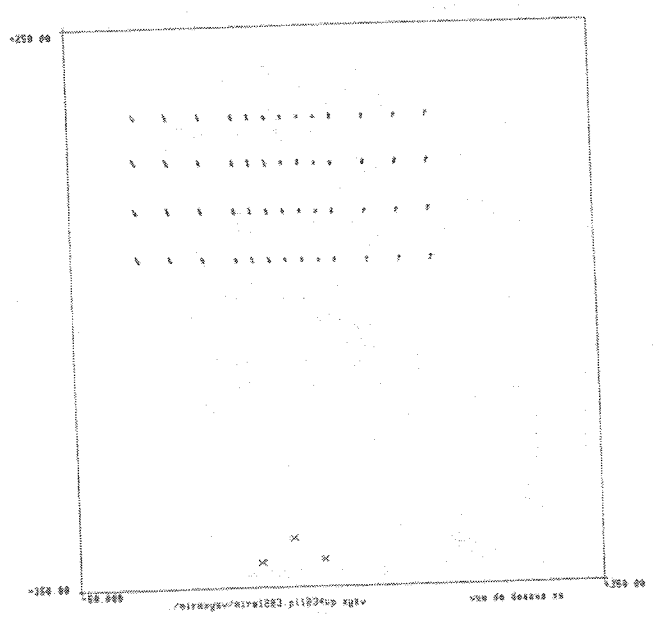
Table 2. Motion between Two Grid Patterns

	r_x (rad)	r_y (rad)	r_z (rad)	t_x (cm)	t_y (cm)	t_z (cm)	Angle (°)
Motion initial estimate	0	0	0	0	0	0	0
Initial covariance	1.0	1.0	1.0	100	100	100	
Motion iteration 1	0.32	0.55	−0.26	−60.70	35.90	20.38	39.82
Motion iteration 2	0.29	0.31	−0.30	−50.90	49.62	49.16	30.27
Real motion manually computed	0.3	0.3	−0.3	−50.00	50.00	50.00	29.77

Fig. 10. Fusion of the reconstructed intersections of the grid pattern.

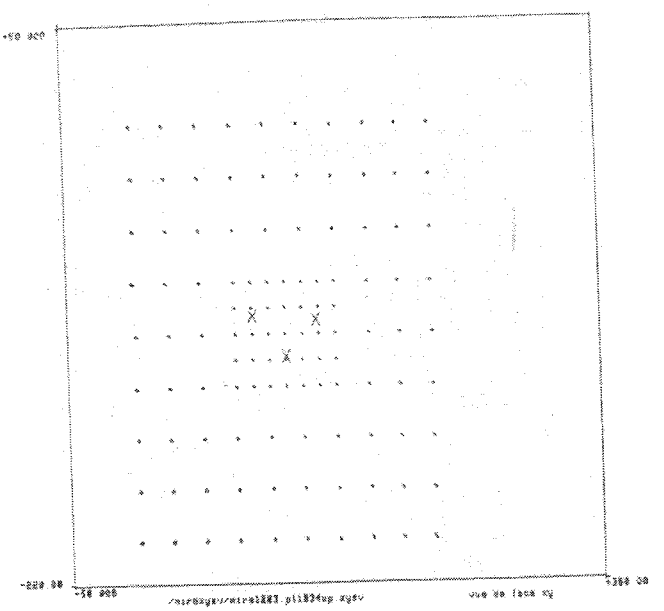


(a)

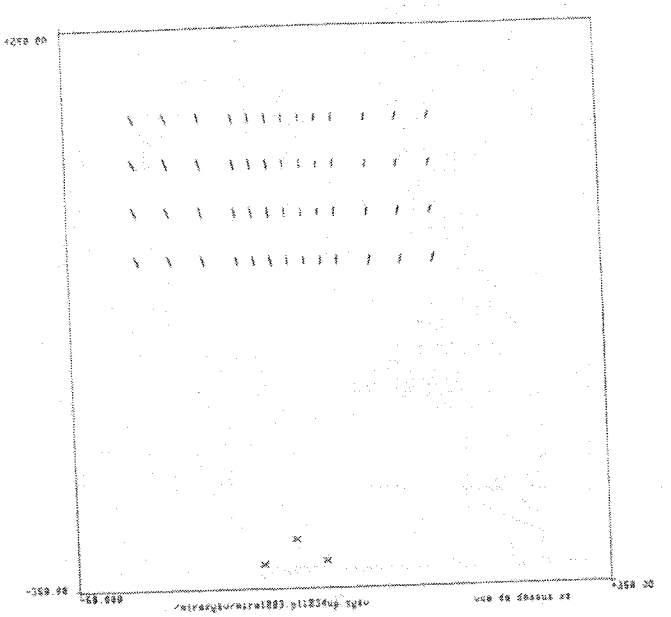


(b)

Fig. 11. Covariance matrices attached to the points shown in Fig. 10.



(a)

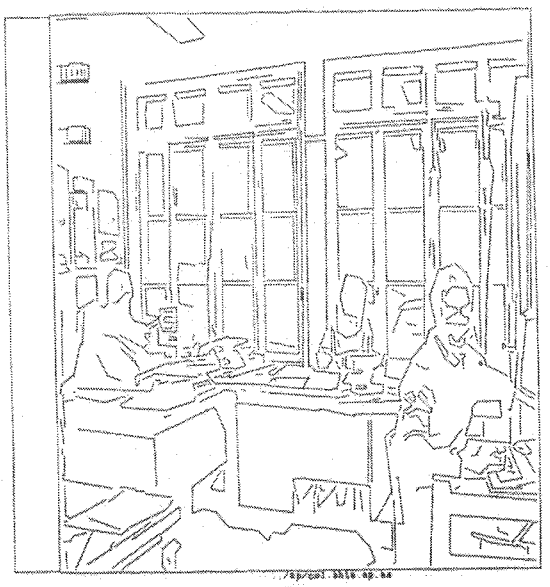


(b)

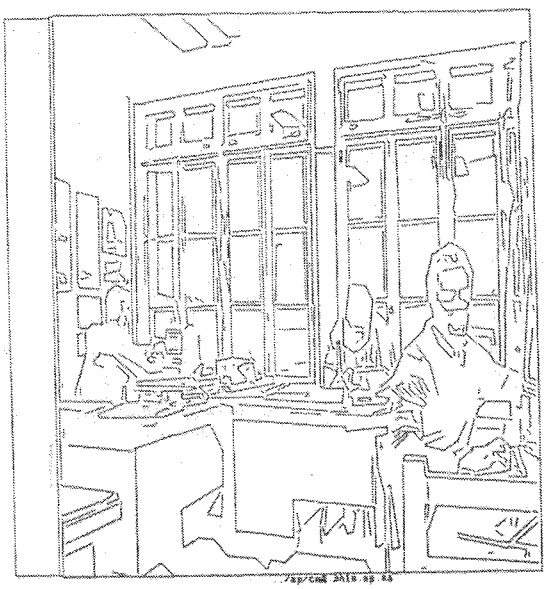
the type (2)). Second, the method of Section 6 has been applied to the points in each of the four coordinate systems associated with views 1 to 4. The results are presented in the same format as the one in Figs. 6 and 7.

Two facts can be inferred from Figs. 10 and 11. First, the uncertainty has been greatly decreased, and, second, the new computed uncertainty still seems to be in excellent agreement with the spread of the updated data, thus yielding more credibility to the whole scheme. In all these figures, the positions of the three cameras are represented by triplets of crosses.

Fig. 12. Polygonal approximation of the edge points of a stereo pair of an office room observed in position 1.

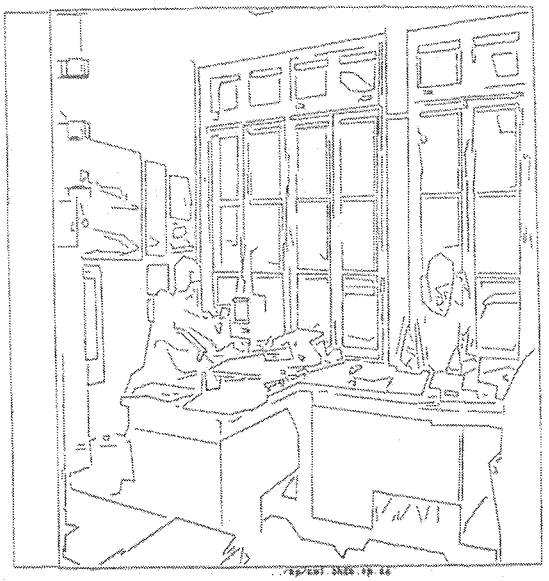


(a)

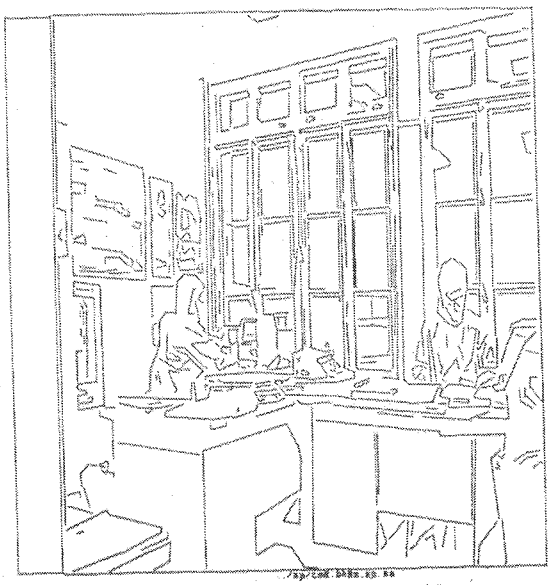


(b)

Fig. 13. Polygonal approximation of the edge points of a stereo pair of the office room in Fig. 12 observed in position 2.



(a)



(b)

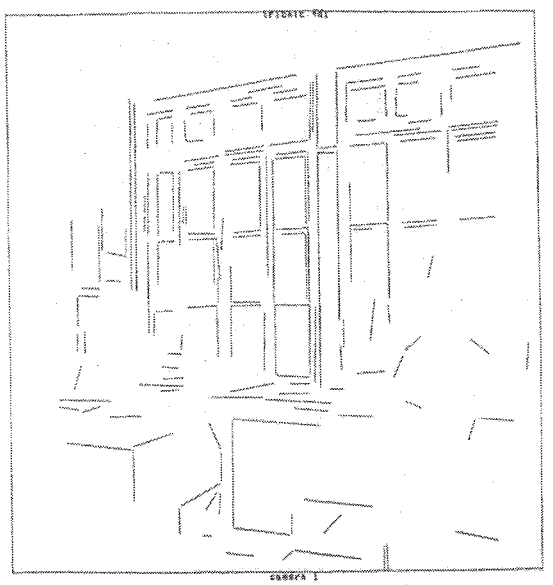
7.2. Office Example

7.2.1. Stereo Reconstruction

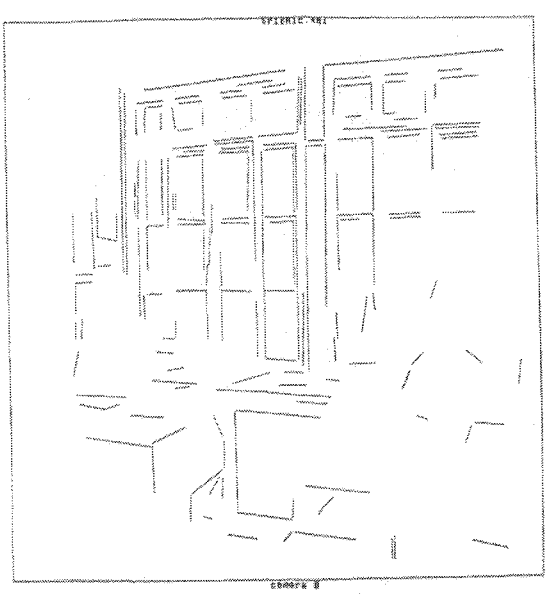
We have applied the same programs to a different set of images. Figures 12 and 13 show the polygonal approximations of the edge points in two stereo pairs of an office scene taken from different viewpoints, and Figs. 14 and 15 show the segments which have been matched by the stereo program described in Ayache and Faverjon (1985).

Figures 16 and 17 have the same format as Fig. 6: i.e., we show the projection of the reconstructed three-dimensional segments in a vertical plane (Figs. 16A

Fig. 14. Edge segments matched in stereo pair of Fig. 12.

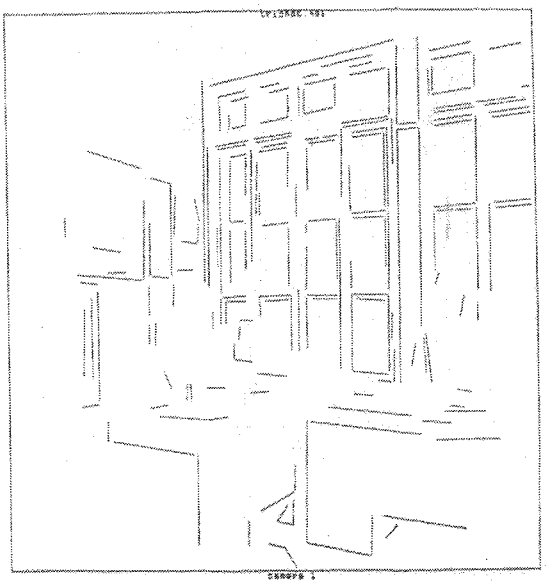


(a)

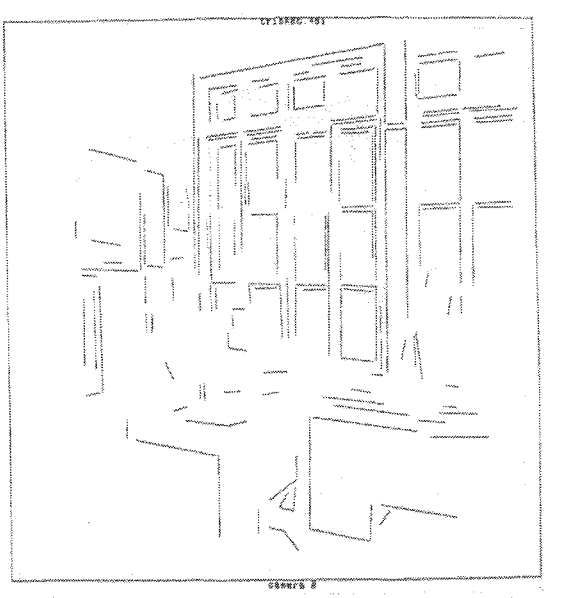


(b)

Fig. 15. Edge segments matched in stereo pair of Fig. 13.



(a)



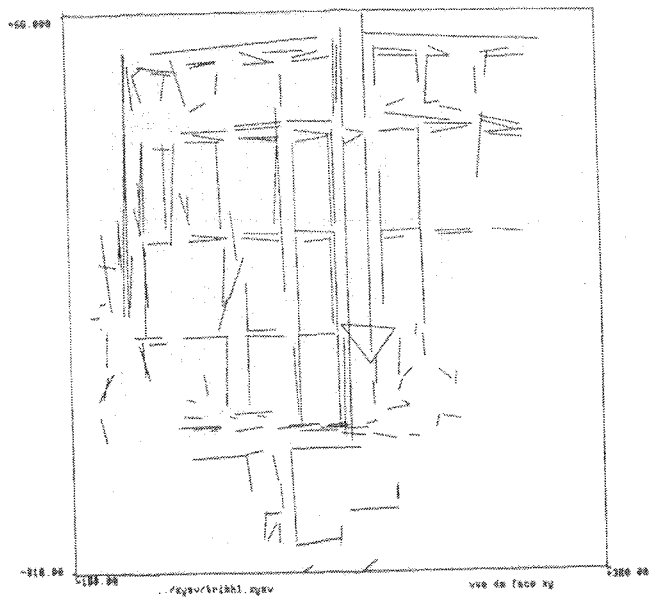
(b)

and 17A) and in a horizontal plane (Figs. 16B and 17B). The position of the cameras mounted on the robot are shown by a triangle. Figures 18 and 19 have the same format as Fig. 7; i.e., we show a representation of the covariance matrices of the endpoints of the reconstructed line segments computed from Eq. (2) in a vertical plane (18A and 19A), and in a horizontal plane (18B and 19B).

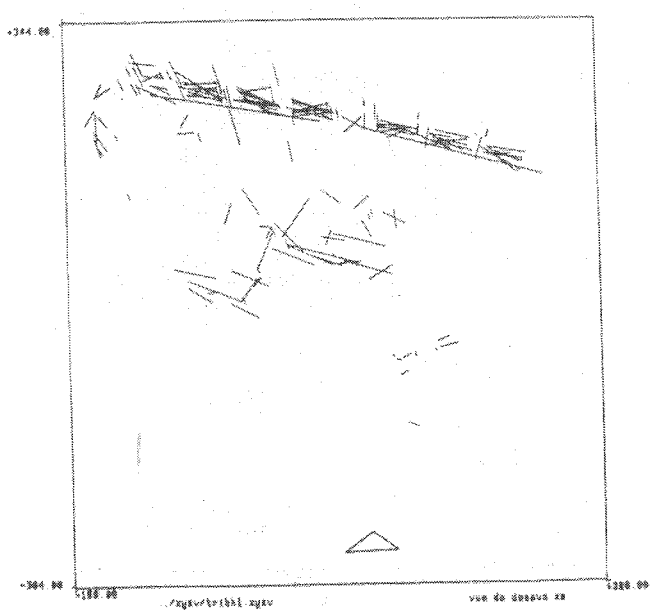
7.2.2. Matching

Figure 20 shows the results of the estimation of the displacement of the robot from view 1 to view 2 by

Fig. 16. Horizontal and vertical projection of the reconstructed segments of the office room observed in position 1.

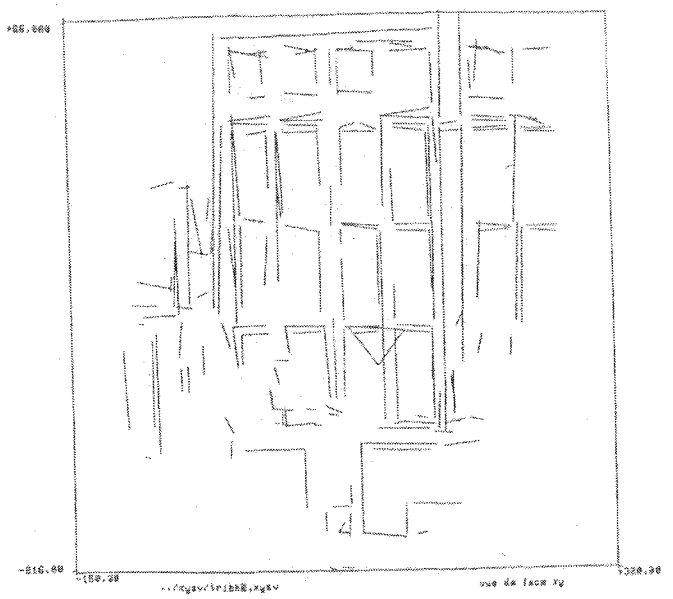


(a)

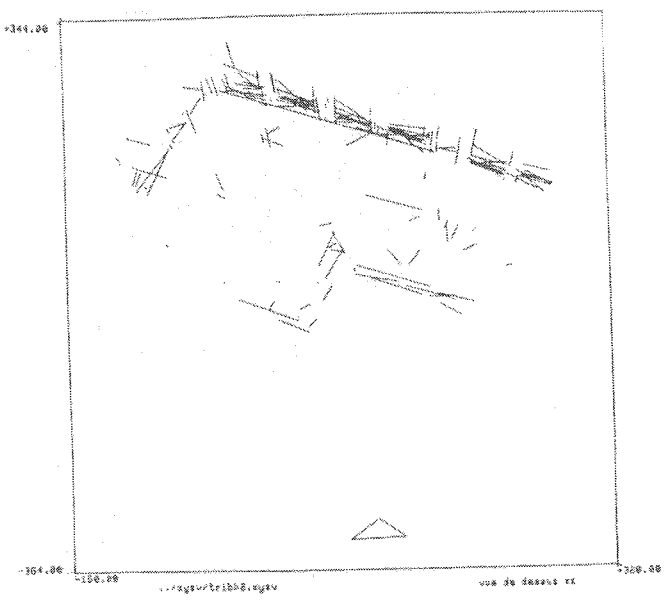


(b)

Fig. 17. Horizontal and vertical projection of the reconstructed segments of the office room observed in position 2.



(a)

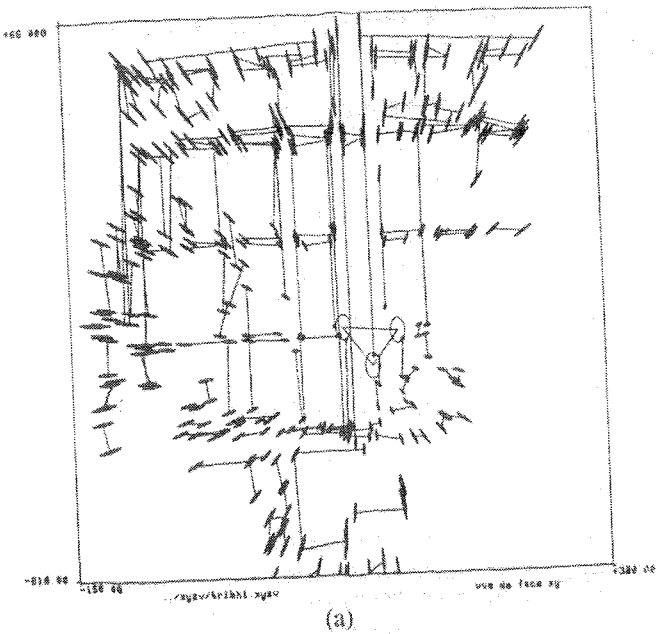


(b)

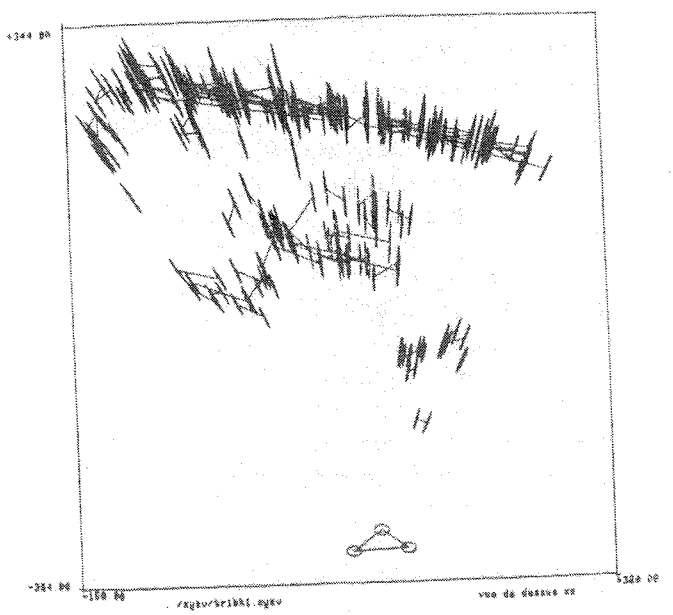
matching three-dimensional lines using Eqs. (4) and (5). Figure 20A shows a view from above of the segments in view 1 which are matched in view 2, Fig. 20B shows the corresponding segments in view 2, and Fig. 21A shows the result of applying the estimated displacement to the segments in view 1. The result looks much like Fig. 20B (as it should). The position of the cameras is shown as a triangle.

Figure 21B is yet another way of displaying the results: After applying to them the estimated displacement, the matched segments in view 1 are projected (as continuous lines) in one of the images corresponding

Fig. 18. Covariance matrices attached to the endpoints of the reconstructed segments of the office room observed in position 1.

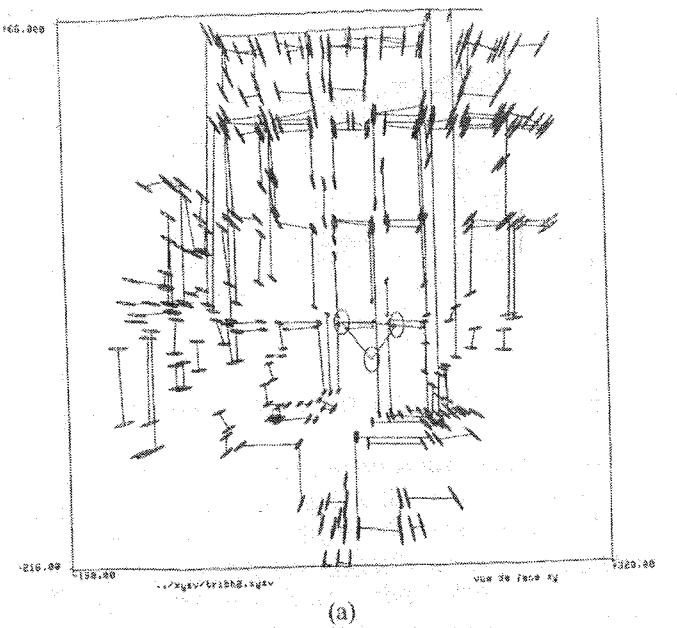


(a)

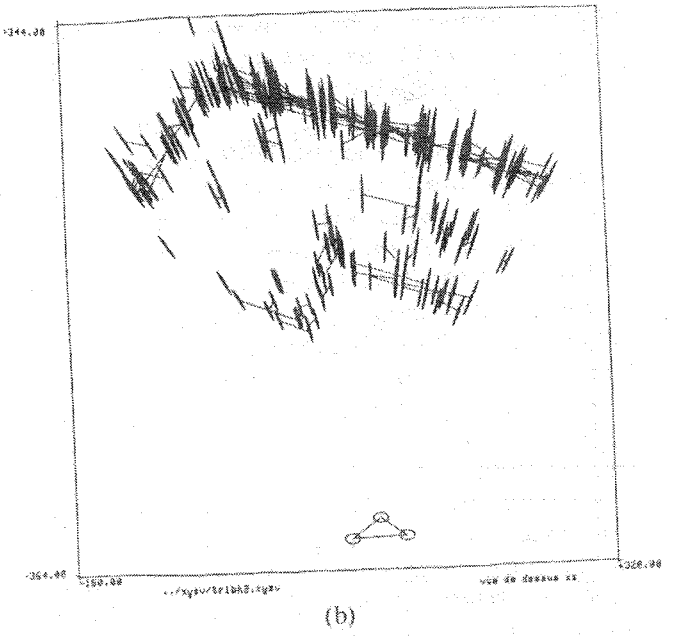


(b)

Fig. 19. Covariance matrices attached to the endpoints of the reconstructed segments of the office room observed in position 2.



(a)



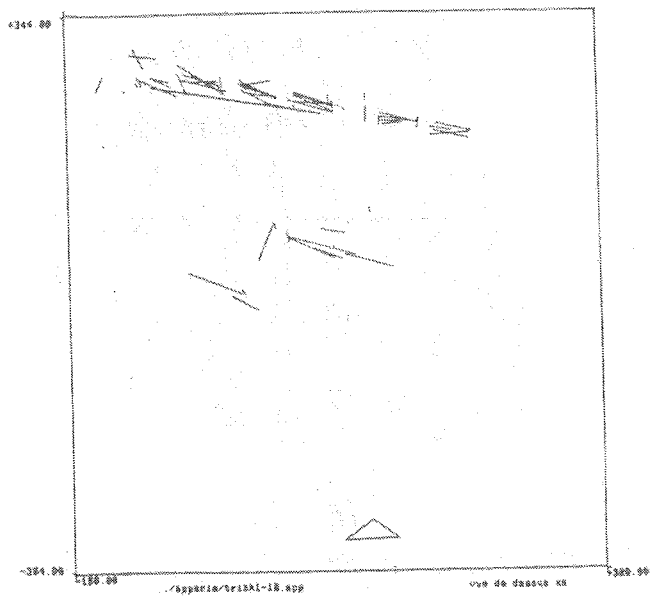
(b)

to view 2 where the segments are displayed as dotted lines. Again, the agreement is seen to be quite good.

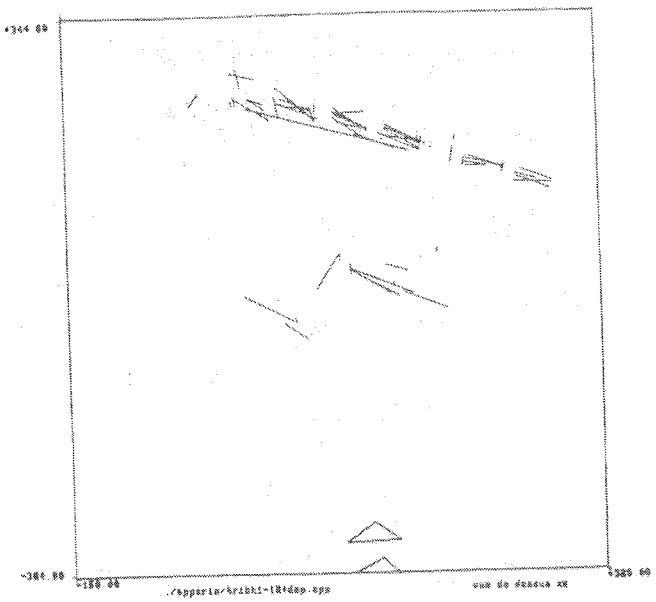
A more quantitative description of what is happening can be found in Table 3. The first row shows the initial estimates of the rotation and translation, the second row the corresponding covariance matrices. These covariance matrices contain the knowledge that the motion is close to a horizontal plane. The third row shows the estimate of the displacement obtained by matching manually two lines in views 1 and 2, thus

Fig. 20. Reconstructed segments matched between positions 1 and 2.

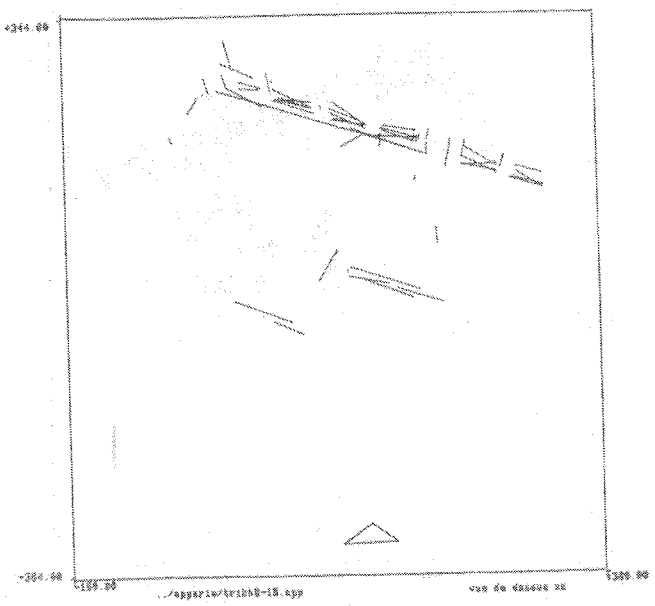
Fig. 21. A. Application of the estimated motion to the segments of position 1: the triangles show the estimated motion of the robot. B. Application of the estimated motion to the segments of position 1 followed by a perspective projection (solid lines) on one of the images actually observed in position 2 (dotted lines).



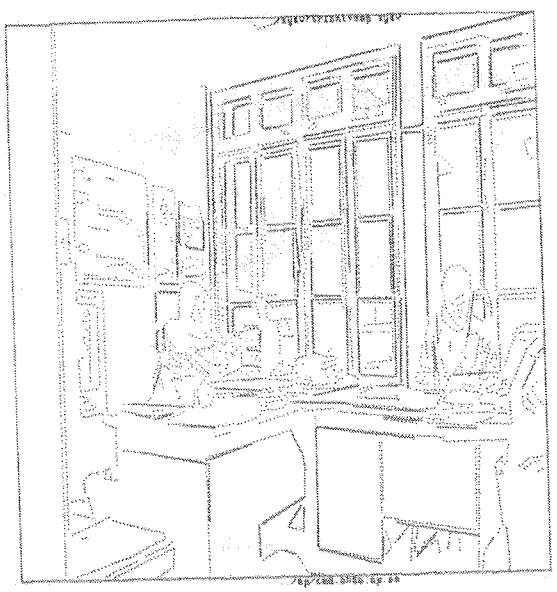
(a)



(a)



(b)



(b)

yielding two measurement equations of types (4) and (5).

The fourth and fifth rows show the estimates found by running the extended Kalman filter twice over a set of 79 and 84 pairs of lines, respectively. Notice that the matches are found automatically by the matcher after the first two matches have been given manually.

7.2.3. Fusion

Experiments concerning fusion of straight lines can be found in Ayache and Faugeras (1987a,b).

Table 3. Motion Computed between Two Stereo Views of an Office

	r_x (rad)	r_y (rad)	r_z (rad)	t_x (cm)	t_y (cm)	t_z (cm)	Angle (°)
Motion initial estimate	0	0	0	0	0	0	0
Initial covariance	0.01	2.0	0.01	500	1	500	
Motion: 2 lines man. matched	0.03	0.17	0.02	42.03	0.32	−24.49	10.17
Motion iteration 1	0.01	0.137	−0.01	47.84	−2.58	−35.01	7.89
Motion iteration 2	0.01	0.138	0.008	47.56	−2.54	−33.56	8.06

8. Conclusions and Discussion

We have proposed some ideas related to the construction, by a mobile robot using stereo, of a three-dimensional model of its environment. In this effort we have been following two guiding lights. The first one is that of linearizing the measurement equations of our processes in order to apply the powerful Kalman filter (extended Kalman filtering). The second one is that of using a representation of three-dimensional rigid displacements which is adapted to such a linearization, and we have come up with the exponential representation of orthogonal matrices. From there, our reflections have pursued three ideas.

The first idea is that the model of the environment must contain not only a geometric characterization of the geometric primitives it uses (here points, lines, and planes), but also a characterization of the uncertainty on the parameters of these primitives, uncertainty caused by the noisy measurements. We have used a characterization of this uncertainty by covariance matrices and shown that it could be traced all the way to the sensor noise (in our case the pixel noise). We call such a description a realistic uncertain description of the environment.

The second idea is that of using local coordinate frames attached to the various positions of the robot to describe the environment and to relate the various representations obtained from several stereo pairs taken from those different viewpoints whose spatial relationships are unknown or imperfectly known, to obtain an estimate of the three-dimensional displacements which relate the positions of the robot and a characterization of the uncertainty of these displacements.

The third idea is to use the previous characterization of their relationships to update the various representations. Indeed, if a given geometric primitive in some coordinate frame has been identified as corresponding to the same part of a physical object than another geometric primitive in another coordinate frame, then we can use our knowledge of the relation between the two frames to update the geometric and uncertainty descriptions of these primitives in their respective coordinate frames.

Several questions remain open. The first one is related to the representation of uncertainty of the geometric primitives such as lines and planes. We have considered that lines were defined by two points. In practice, this is not so since three-dimensional lines are reconstructed by intersecting planes defined by the focal center of the cameras and two-dimensional line segments obtained by mean square approximation of edge points. Therefore the uncertainty on three-dimensional lines is more complex than the uncertainty on pairs of three-dimensional points. Second, the fact that we are assuming Gaussian distributions on the parameters of the representations of lines, planes, and rotations is also subject to criticism for two reasons. The first one is unavoidable and is related to the linearization of the measurement equations; the second one is deeper and related to the fact that we have to be careful when defining probability densities on geometric sets. Despite these problems, the results that we have obtained are quite encouraging, and we think that the basic approach is promising and worth more investigation. New developments and results are presented in Ayache and Faugeras (1987,1988).

Acknowledgments

The authors acknowledge the contribution of Bernard Faverjon to an earlier version of this work (Faugeras et al. 1986) and stimulating discussions with Francis Lustman, Giorgio Toscani, and Pierre Tournassoud. Also, the help of Chantal Chazelas and Nathalie Rocher in preparing this article is much appreciated.

Appendix A: Kalman Filter Equations

The equation

$$\mathbf{y}_i = \mathbf{M}_i\mathbf{a} + \mathbf{u}_i$$

of Section 2 is, in the terminology of the Kalman filter (Jazwinsky 1970), a measurement equation on the process **a** constant with respect to i: $\mathbf{a}_i = \mathbf{a}_{i-1}$. Using the equations of the Kalman filter, we write the new estimate and estimation covariance matrix as

$$\begin{aligned}\hat{\mathbf{a}}_i &= \hat{\mathbf{a}}_{i-1} + \mathbf{K}_i(\mathbf{y}_i - \mathbf{M}_i\hat{\mathbf{a}}_{i-1}),\\ \mathbf{K}_i &= \mathbf{S}_{i-1}\mathbf{M}_i^t(\mathbf{W}_i + \mathbf{M}_i\mathbf{S}_{i-1}\mathbf{M}_i^t)^{-1},\\ \mathbf{S}_i &= (\mathbf{I} - \mathbf{K}_i\mathbf{M}_i)\mathbf{S}_{i-1},\end{aligned}$$

or, equivalently,

$$\mathbf{S}_i^{-1} = \mathbf{S}_{i-1}^{-1} + \mathbf{M}_i^t\mathbf{W}_i^{-1}\mathbf{M}_i,$$

where

$$\mathbf{W}_i = E(\mathbf{u}_i\mathbf{u}_i^t).$$

When all measurements have been processed, parameter **a** is known by its a posteriori estimate $\hat{\mathbf{a}}_n$ and the corresponding covariance matrix $\mathbf{S}_n$.

Appendix B: Stereo Reconstruction

Remembering that $\mathbf{x} = [u, v]^t$ and $\mathbf{a} = \mathbf{OM}$, we can easily derive from Eq. (2) that

$$\frac{\partial \mathbf{f}}{\partial \mathbf{a}} = \begin{bmatrix} u\mathbf{l}_3 - \mathbf{l}_1 \\ v\mathbf{l}_3 - \mathbf{l}_2 \end{bmatrix} \quad \text{a } 2 \times 3 \text{ matrix}$$

and

$$\frac{\partial \mathbf{f}}{\partial \mathbf{x}} = \begin{bmatrix} \mathbf{l}_3\mathbf{OM} + l_{34} & 0 \\ 0 & \mathbf{l}_3\mathbf{OM} + l_{34} \end{bmatrix}.$$

Appendix C: Exponential Representation of Rotation Matrices

A simple justification of the fact that every matrix $e^{\mathbf{H}}$, with **H** an antisymmetric matrix, is orthogonal, is the following:

$$(e^{\mathbf{H}})^t = e^{\mathbf{H}^t} = e^{-\mathbf{H}} = (e^{\mathbf{H}})^{-1}.$$

This follows from the formula $e^{\mathbf{H}} = \mathbf{I} + \mathbf{H}/1! + \mathbf{H}^2/2! + \cdots$. We also know (Gantmacher 1977) that the eigenvalues a and b of **R** and **H** are related by $a = e^b$. Since it is well known that the eigenvalues of **R** are 1, $e^{i\theta}$, and $e^{-i\theta}$, where θ is the rotation angle, the eigenvalues of **H** are easily found to be 0, $i(a^2 + b^2 + c^2)^{1/2}$, and $-i(a^2 + b^2 + c^2)^{1/2}$. Therefore

$$\theta = \pm(a^2 + b^2 + c^2)^{1/2}.$$

Let us now consider an eigenvector of **H** associated with eigenvalue 0. Since **H** represents the vector product with vector $\mathbf{v} = [a, b, c]^t$, **v** is such a vector:

$$\mathbf{Hv} = \mathbf{0}.$$

v is also an eigenvector of matrix **R** associated with eigenvalue 1, as can be easily verified by the formula

$$\mathbf{R} = \mathbf{e}^{\mathbf{H}} = \mathbf{I} + \mathbf{H}/1! + \mathbf{H}^2/2! + \cdots.$$

Therefore

$$\mathbf{Rv} = \mathbf{v},$$

and **v** gives the direction of the axis of rotation. **H** is

thus a very compact representation of the rotation in terms of its axis and angle. Another property of this representation can be deduced from a theorem in Gantmacher (1977, p. 158), which states that if we compute the Lagrange-Sylvester polynomial p of the exponential function for the eigenvalues of $\mathbf{H}$—i.e., the polynomial such that

$$p(0) = e^0 = 1, \quad p(i\theta) = e^{i\theta}, \quad p(-i\theta) = e^{-i\theta},$$

—then we have the nice relationship

$$e^{\mathbf{H}} = p(\mathbf{H}).$$

It can be easily verified that

$$p(\mathbf{H}) = \mathbf{I} + ((\sin\theta)/\theta)\mathbf{H} + ((1 - \cos\theta)/\theta^2)\mathbf{H}^2,$$

which is precisely the well-known Rodrigues formula (Rodrigues 1840).

We finish with a property of this representation that is heavily used in the rest of the paper and related to the derivative of $e^{\mathbf{H}}$ with respect to $\mathbf{r}$. The corresponding computation is a little painful but is certainly worth doing. We refer the interested reader to the INRIA internal report corresponding to this paper or to Ayache (1988). We give only the result. We want to compute the 3×3 matrix

$$\mathbf{K}(\mathbf{R}, \mathbf{OM}) = \frac{\partial(\mathbf{ROM})}{\partial \mathbf{r}}$$

or, more simply, the 3-vector $\mathbf{K}(\mathbf{R}, \mathbf{OM})\mathbf{v}$, where $\mathbf{v}$ is an arbitrary 3-vector.

Letting $f(\theta) = (\sin\theta)/\theta$ and $g(\theta) = (1 - \cos\theta)/\theta^2$, one obtains

$$\begin{aligned}\mathbf{K}(\mathbf{R}, \mathbf{OM})\mathbf{v} = {} & f'(\theta)/\theta(\mathbf{r} \cdot \mathbf{v})\mathbf{r} \times \mathbf{OM} \\ & + g'(\theta)/\theta(\mathbf{r} \cdot \mathbf{v})\mathbf{r} \times (\mathbf{r} \times \mathbf{OM}) \\ & + f(\theta)(\mathbf{v} \times \mathbf{OM}) \\ & + g(\theta)(\mathbf{v} \times (\mathbf{r} \times \mathbf{OM}) \\ & + \mathbf{r} \times (\mathbf{v} \times \mathbf{OM})),\end{aligned} \tag{C.1}$$

where $f'(\theta)$ and $g'(\theta)$ are the derivatives of $f(\theta)$ and $g(\theta)$ with respect to θ. (A slightly simpler version of this equation will be found in Ayache 1988.) A special case occurs when $\mathbf{v}$ is collinear to $\mathbf{r}$ (i.e., when $\mathbf{v} = \mathbf{r}$). In this case one has

$$\mathbf{K}(\mathbf{R}, \mathbf{OM})\mathbf{v} = \alpha \mathbf{r} \times \mathbf{ROM}.$$

Appendix D: Matching Three-Dimensional Points

Using Eq. (3), remembering that $\mathbf{x} = [\mathbf{O'M'}^t, \mathbf{OM}^t]^t$, and $\mathbf{a} = [\mathbf{r}^t, \mathbf{t}^t]^t$, and taking into account Eq. (C.1), we have, with the notations of Section 1,

$$\partial \mathbf{f}/\partial \mathbf{x} = [\mathbf{I} \quad -\mathbf{R}] \qquad \text{a } 3 \times 6 \text{ matrix}$$

and

$$\partial \mathbf{f}/\partial \mathbf{a} = [-\mathbf{K}(\mathbf{R}, \mathbf{OM}) \quad -\mathbf{I}] \qquad \text{a } 3 \times 6 \text{ matrix.}$$

Appendix E: Matching Three-Dimensional Lines

We use Eqs. (4) and (5). From these equations, we can take $\mathbf{x} = [\mathbf{l'}^t, \mathbf{m'}^t, \mathbf{l}^t, \mathbf{m}^t]^t$ and, as in the case of points, $\mathbf{a} = [\mathbf{r}^t, \mathbf{t}^t]^t$. From these, we can deduce

$$\frac{\partial \mathbf{f}}{\partial \mathbf{l'}} = \begin{bmatrix} -c(\mathbf{Rl}) \\ -c(\mathbf{m'} - \mathbf{Rm} - \mathbf{t}) \end{bmatrix},$$

$$\frac{\partial \mathbf{f}}{\partial \mathbf{m'}} = \begin{bmatrix} \mathbf{0} \\ c(\mathbf{l'}) \end{bmatrix},$$

$$\frac{\partial \mathbf{f}}{\partial \mathbf{l}} = \begin{bmatrix} c(\mathbf{l'})\mathbf{R} \\ \mathbf{0} \end{bmatrix},$$

$$\frac{\partial \mathbf{f}}{\partial \mathbf{m}} = \begin{bmatrix} \mathbf{0} \\ -c(\mathbf{l'})\mathbf{R} \end{bmatrix},$$

which completes the computation of $\partial \mathbf{f}/\partial \mathbf{x} = [\partial \mathbf{f}/\partial \mathbf{m'}, \partial \mathbf{f}/\partial \mathbf{l'}, \partial \mathbf{f}/\partial \mathbf{m}, \partial \mathbf{f}/\partial \mathbf{l}]$, a 6×12 matrix. We also have

$$\frac{\partial \mathbf{f}}{\partial \mathbf{r}} = \begin{bmatrix} c(\mathbf{l'})\mathbf{K}(\mathbf{R}, \mathbf{l}) \\ -c(\mathbf{l'})\mathbf{K}(\mathbf{R}, \mathbf{m}) \end{bmatrix},$$

$$\frac{\partial \mathbf{f}}{\partial \mathbf{t}} = \begin{bmatrix} \mathbf{0} \\ -c(\mathbf{l'}) \end{bmatrix},$$

which completes the computation of $\partial \mathbf{f}/\partial \mathbf{a} = [\partial \mathbf{f}/\partial \mathbf{r}, \partial \mathbf{f}/\partial \mathbf{t}]$, a 6×6 matrix. But, as we said before, only four measurement equations can be kept instead of six; therefore, one must keep for $\partial \mathbf{f}/\partial \mathbf{x}$ (and $\partial \mathbf{f}/\partial \mathbf{a}$) only the 4×12 (resp. 4×6) matrix obtained by keeping the four rows corresponding to the four independent measurements.

Appendix F: Matching Planes

We use Eqs. (6) and (7). Remembering that $\mathbf{x} = [\mathbf{n}'^{\mathrm{t}}, \mathbf{n}^{\mathrm{t}}, d', d]^{\mathrm{t}}$ and $\mathbf{a} = [\mathbf{r}^{\mathrm{t}}, \mathbf{t}^{\mathrm{t}}]^{\mathrm{t}}$ and taking into account Eq. (C.1), we have, with the notations of Section 2,

$$\frac{\partial \mathbf{f}}{\partial \mathbf{x}} = \begin{bmatrix} \mathbf{I} & -\mathbf{R} & 0 & 0 \\ -\mathbf{t}^{\mathrm{t}} & \mathbf{0}^{\mathrm{t}} & -1 & 1 \end{bmatrix} \quad \text{a } 4 \times 8 \text{ matrix}$$

and

$$\frac{\partial \mathbf{f}}{\partial \mathbf{a}} = \begin{bmatrix} -\mathbf{K}(\mathbf{R}, \mathbf{n}) & \mathbf{0} \\ \mathbf{0}^{\mathrm{t}} & -\mathbf{n}'^{\mathrm{t}} \end{bmatrix} \quad \text{a } 4 \times 6 \text{ matrix,}$$

where $\mathbf{0}$ is a 3×1 vector of zeros.

Appendix G: Matching Three-Dimensional and Two-Dimensional Points and Lines

We treat the case of points first. From Eq. (8), and remembering that $\mathbf{x} = [\mathbf{OM}^{\mathrm{t}}, u', v']^{\mathrm{t}}$ and $\mathbf{a} = [\mathbf{r}^{\mathrm{t}}, \mathbf{t}^{\mathrm{t}}]^{\mathrm{t}}$, we have

$$\frac{\partial \mathbf{f}}{\partial \mathbf{a}} = \begin{bmatrix} (u'\mathbf{l}_3 - \mathbf{l}_1)\mathbf{K}(\mathbf{R}, \mathbf{OM}) & u'\mathbf{l}_3 - \mathbf{l}_1 \\ (v'\mathbf{l}_3 - \mathbf{l}_2)\mathbf{K}(\mathbf{R}, \mathbf{OM}) & v'\mathbf{l}_3 - \mathbf{l}_2 \end{bmatrix} \quad \text{a } 2 \times 6 \text{ matrix,}$$

$$\frac{\partial \mathbf{f}}{\partial \mathbf{x}} = \begin{bmatrix} (u'\mathbf{l}_3 - \mathbf{l}_1)\mathbf{R} & \mathbf{l}_3(\mathbf{ROM} + \mathbf{t}) + l_{34} & 0 \\ (v'\mathbf{l}_3 - \mathbf{l}_2)\mathbf{R} & 0 & \mathbf{l}_3(\mathbf{ROM} + \mathbf{t}) + l_{34} \end{bmatrix}$$

a 2×5 matrix.

Let us now tackle the case of lines. Remembering that $\mathbf{x} = [\mathbf{OM}_1^{\mathrm{t}}, \mathbf{OM}_2^{\mathrm{t}}, u_1, v_1, u_2, v_2]^{\mathrm{t}}$ and $\mathbf{a} = [\mathbf{r}^{\mathrm{t}}, \mathbf{t}^{\mathrm{t}}]^{\mathrm{t}}$, we can compute the components of $\mathbf{h}_1$:

$$\begin{aligned} h_{11} &= (\mathbf{l}_2(\mathbf{ROM}_1 + \mathbf{t}) + l_{24})(\mathbf{l}_3(\mathbf{ROM}_2 + \mathbf{t}) + l_{34}) \\ &\quad - (\mathbf{l}_2(\mathbf{ROM}_2 + \mathbf{t}) + l_{24})(\mathbf{l}_3(\mathbf{ROM}_1 + \mathbf{t}) + l_{34}), \\ h_{12} &= (\mathbf{l}_1(\mathbf{ROM}_2 + \mathbf{t}) + l_{14})(\mathbf{l}_3(\mathbf{ROM}_1 + \mathbf{t}) + l_{34}) \\ &\quad - (\mathbf{l}_1(\mathbf{ROM}_1 + \mathbf{t}) + l_{14})(\mathbf{l}_3(\mathbf{ROM}_2 + \mathbf{t}) + l_{34}), \\ h_{13} &= (\mathbf{l}_1(\mathbf{ROM}_1 + \mathbf{t}) + l_{14})(\mathbf{l}_2(\mathbf{ROM}_2 + \mathbf{t}) + l_{24}) \\ &\quad - (\mathbf{l}_1(\mathbf{ROM}_2 + \mathbf{t}) + l_{14})(\mathbf{l}_2(\mathbf{ROM}_1 + \mathbf{t}) + l_{24}). \end{aligned}$$

Using standard rules of differential calculus, we can write

$$\frac{\partial \mathbf{f}}{\partial \mathbf{x}} = \left[\frac{-c(\mathbf{h}_2)\partial \mathbf{h}_1}{\partial \mathbf{OM}_1} \quad \frac{-c(\mathbf{h}_2)\partial \mathbf{h}_1}{\partial \mathbf{OM}_2} \quad \frac{c(\mathbf{h}_1)\partial \mathbf{h}_2}{\partial [u_1, v_1]} \quad \frac{c(\mathbf{h}_1)\partial \mathbf{h}_2}{\partial [u_2, v_2]} \right]^{\mathrm{t}},$$

which is a 3×10 matrix, and

$$\frac{\partial \mathbf{f}}{\partial \mathbf{a}} = \left[\frac{-c(\mathbf{h}_2)\partial \mathbf{h}_1}{\partial \mathbf{r}} \quad \frac{-c(\mathbf{h}_2)\partial \mathbf{h}_1}{\partial \mathbf{t}} \right]^{\mathrm{t}},$$

which is a 3×6 matrix. The various partial derivatives can be easily computed by using the expressions for $\mathbf{h}_1$ and $\mathbf{h}_2$ and Eq. (C.1). Of course, as in the case of line matching, one eventually keeps for $\partial \mathbf{f}/\partial \mathbf{x}$ (resp. $\partial \mathbf{f}/\partial \mathbf{a}$) the 2×10 matrix (resp. 2×6) obtained by keeping the two rows corresponding to the two independent measurements.

References

Ayache, N. 1988. Construction et fusion de représentations visuelles tridimensionnelles—applications à la robotique mobile. Thèse d'Etat.

Ayache, N., and Faugeras, O. D. 1986. Hyper: a new approach for the recognition and positioning of two-dimensional objects. *IEEE Trans. PAMI* PAMI-8(1):44–54.

Ayache, N., and Faugeras, O. D. 1987 (Aug., Milano, Italy). Building a consistent 3D representation of a mobile robot environment by combining multiple stereo views. *Proc. Int. Joint. Conf. on Artificial Intelligence,* pp. 808–810.

Ayache, N., and Faugeras, O. D. 1988. Maintaining representations of the environment of a mobile robot. *Robotics Research, the Fourth Int. Symp.* Cambridge, Mass: MIT Press, pp. 337–350.

Ayache, N., Faugeras, O. D., Faverjon, B., and Toscani, G. 1985 (October 13–16, Bellaire, TX). Matching depth maps obtained by passive stereo. *Proc. Third Workshop*

on Computer Vision: Representation and Control, pp. 197–204.

Ayache, N., and Faverjon, B. 1987. Efficient registration of stereo images by matching graph descriptions of edge segments. *Int. J. Computer Vision,* pp. 107–131.

Bolle, R. M., and Cooper, D. B. 1985 (Dec., Cannes, France). On parallel bayesian estimation and recognition for large data sets, with application to estimating 3-D complex-object position from range data. *SPIE Conf. on Vision for Robots.*

Brooks, R. 1985. Aspects of mobile robot visual map making. In *Robotics Research, the Second Int. Symp.* Cambridge, Mass.: MIT Press, pp. 369–376.

Crowley, J. L. 1986 (April 7–10, San Francisco, Calif.). Representation and maintenance of a composite surface model. *IEEE Conf. on Robotics and Automation,* pp. 1455–1462.

Darmon, C. A. 1982 (Paris, France). A recursive method to apply the Hough transform to a set of moving objects. *Proc. ICASSP 82,* pp. 825–829.

Durrant-Whyte, H. F. 1986 (April 7–10, San Francisco, Calif.). Consistent integration and propagation of disparate sensor observations. *Proc. IEEE Conf. on Robotics and Automation,* pp. 1464–1469.

Faugeras, O. D., and Toscani, G. 1986 (Miami Beach, Fl.). The calibration problem for stereo. *Proc. CVPR86,* pp. 15–20.

Faugeras, O. D., Ayache, N., and Faverjon, B. 1986 (April 7–10, San Francisco, Calif.). Building visual maps by combining noisy stereo measurements. *Proc. IEEE Conf. on Robotics and Automation,* pp. 1433–1438.

Faugeras, O. D., and Hebert, M. 1986. The representation, recognition, and localisation of 3D shapes from range data. *Int. J. Robotics Res.* 5(3):27–52.

Faugeras, O. D., and Lustman, F. 1986 (July 21–25, Brighton Centre, Great Britain). Inferring planes by hypothesis prediction and verification for a mobile robot. *European Conf. on Artificial Intelligence,* pp. 143–147.

Gantmacher, F. R. 1977. *Matrix theory.* New York: Chelsea.

Jazwinsky, A. M. 1970. *Stochastic processes and filtering theory.* New York: Academic Press.

Laumond, J. P., and Chatila, R. 1985 (Saint Louis, Mo.). Position referencing and consistent world modeling for mobile robots. *Proc. IEEE Conf. on Robotics and Automation,* pp. 138–145.

Lowe, D. 1985. *Perceptual organization and visual recognition.* Boston: Kluwer Academic.

Rodrigues, O. 1840. Des lois géométriques qui régissent les déplacements d'un système solide dans l'espace, et de la variation des coordonnées provenant de ces déplacements considérés indépendamment des causes qui peuvent les produire. *J. Math. Pures Appl.* 5(1):380–440.

Smith, R. C., and Cheeseman, P. 1986. On the representation and estimation of spatial uncertainty. *Int. J. Robotics Res.* 5(4):56–68.

Stuelpnagel, J. 1964. On the parameterization of the three-dimensional rotation group. *SIAM Rev.* 6(4):422–430.