Open access • Journal Article • DOI:10.1177/027836498800700605

Building, registrating, and fusing noisy visual maps — Source link

Nicholas Ayache, Olivier Faugeras

Institutions: French Institute for Research in Computer Science and Automation

Published on: 01 Dec 1988 - The International Journal of Robotics Research (SAGE Publications)

Topics: Geometric primitive and Mobile robot

Building,registrating,and fusing noisy visual maps

Nicholas Ayache, Olivier Faugeras

INRIA

CENTRE DE ROCQUENCOURT

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P.105
78153 Le Chesnay Cedex
France
Tél.: (1) 39 63 55 11

Rapports de Recherche

N° 596

BUILDING, REGISTRATING, AND FUSING NOISY VISUAL MAPS

Nicholas AYACHE
Olivier D. FAUGERAS

Décembre 1986

Building, Registrating, and Fusing Noisy Visual Maps

Nicholas Ayache and Olivier D. Faugeras

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

ABSTRACT

This paper deals with the problem of building 3D 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 both information about the geometry of the environment and information 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. Finally 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, i.e. building Visual Maps, registrating and fusing them, within the general framework of Extended Kalman Filtering which allows to efficiently combine measurements in the presence of noise.

Keywords

3D Descriptions, Visual Maps, Registration, Fusion, Uncertainty, Stereo, Motion, Mobile Robot, Extended Kalman Filtering, 3D Rotations.

Construction, Appariement et Mise à Jour de Cartes Visuelles Tridimensionnelles Bruitées

RESUME

Cet article traite du problème de la construction de représentations 3D de l'environnement d'un robot mobile en utilisant la Vision passive. Nous appelons ces représentations des cartes visuelles. Elles sont locales, c'est-à-dire rapportées à des sytèmes de coordonnées locaux.

Du fait de la présence de bruit, elles incorporent à la fois de l'information sur la géométrie de l'environnement mais aussi sur l'incertitude des paramètres définissant cette géométrie. L'incertitude géométrique est quantitativement reliée à sa source c'est-à-dire le bruit des capteurs.

Nous montrons comment mettre en correspondance des cartes visuelles construites dans différentes positions du robot afin de calculer une meilleure estimée de son déplacement, en supposant que l'environnement est statique.

Enfin, nous utilisons ces estimées pour fusionner les différentes cartes visuelles et réduire localement l'incertitude des primitives géométriques qui ont trouvé des correspondants dans d'autres cartes.

Nous proposons d'effectuer ces trois tâches, la construction des cartes visuelles, leur mise en correspondance, et leur fusion, à l'aide du filtre de Kalman étendu qui permet de combiner de manière efficace des mesures bruitées.

Mots Clés

Descriptions 3D, Cartes visuelles, Mise en correspondance, Fusion, Incertitude, Stéréo, Mouvement, Robot mobile, Filtre de Kalman étendu, Rotations 3D.

1. Introduction

The problem of dealing with noise in 3D 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 explicitely 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 3D information. This work continues the one presented in (Ayache-Faugeras-Faverjon 1985 and Faugeras-Ayache-Faverjon 1986) and is connected to that of Brooks (Brooks 1985), Bolle and Cooper (Bolle and Cooper 1985), Laumond and Chatila (Laumond and Chatila 1985), Crowley (Crowley 1986), Smith and Cheeseman (Smith and Cheeseman 1986), and Durrant-White (Durrant-White 1986).

The goals to be reached are the establishment of a 3-D 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, navigation strategies : 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 3D rotations to relate frames of reference, as matrix exponentials of antisymmetric matrixes.

2. Linearizing the problem

In all the cases we discuss, we deal with an observation $\mathbf{x}$ in $\mathbf{IR}^m$ that depends on a parameter $\mathbf{a}$ in $\mathbf{IR}^n$ in a nonlinear fashion that can be expressed as a relation

$$\mathbf{f}(\mathbf{x},\mathbf{a}) = \mathbf{0}$$

where $\mathbf{f}$ maps $\mathbf{IR}^m \times \mathbf{IR}^n$ into $\mathbf{IR}^p$. We assume that the observation $\mathbf{x}$ is corrupted with noise which we model as an additive zero mean Gaussian noise :

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

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

Supposing that we know a "good" estimate $\mathbf{a}^*$ of $\mathbf{a}$, then we can use the idea of linearization and expand $\mathbf{f}$ in the vicinity $(\mathbf{x}',\mathbf{a})$:

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

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

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

which is a linear measurement equation :

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

where :

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

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

Notice that **y** and **M** are known and **u**'s second order statisitics are also known :

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

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

We "measure" **y** because we know **f** and **a***, and observe **x**, the corrupted version of **x'**. For the same reasons, matrixes **M** and $\partial \mathbf{f}(\mathbf{x},\mathbf{a^*})/\partial \mathbf{x}$ are known. So, if we have n measurements $\mathbf{x}_1,\ldots,\mathbf{x}_n$ satisfying $\mathbf{f}_i(\mathbf{x}_i,\mathbf{a})=\mathbf{0}$ and if we start with an initial estimate $\hat{\mathbf{a}}_0$ of **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 **a** and its covariance matrix $\mathbf{S}_n=E((\hat{\mathbf{a}}_n-\mathbf{a})(\hat{\mathbf{a}}_n-\mathbf{a})^t)$ 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:

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

This equation is important, because it shows how the Kalman filtering explicitely takes into account the 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 details the application of this technique to a number of the previously mentioned problems.

3. Stereo matching

In a standard, two cameras, stereo system such as the one depicted in figure 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:

$$\begin{bmatrix} s_i\ u_i \\ s_i\ v_i \\ s_i \end{bmatrix} = P_i \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \qquad i = 1,2$$

where matrices 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 3-D 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 3x4 matrix **P** as :

$$\mathbf{P} = \begin{bmatrix} l_1 & l_{14} \\ l_2 & l_{24} \\ l_3 & l_{34} \end{bmatrix}$$

where l_1,l_2,l_3 are 3x1 row vectors, and eliminating s, we have the following relationships:

$$(l_3\,\mathbf{OM} + l_{34})u - l_1\,\mathbf{OM} - l_{14} = 0$$
$$(l_3\,\mathbf{OM} + l_{34})v - l_2\,\mathbf{OM} - l_{24} = 0$$

and therefore:

$$\mathbf{f}(\mathbf{x},\mathbf{a}) = \begin{bmatrix} (l_3\mathbf{OM} + l_{34})u - l_1\mathbf{OM} - l_{14} \\ (l_3\,\mathbf{OM} + l_{34})v - l_2\,\mathbf{OM} - l_{24} \end{bmatrix} \qquad (2)$$

From the above formalism we can deduce an estimate $\hat{\mathbf{a}}_2$ of **a** after two measurements i.e. of the position in 3-D 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 S_0. The equations are in Appendix B. In practice, S_0 is diagonal with very large diagonal terms, so that the first term in equation (1) is very small.

This shows how the uncertainty on points obtained by stereo can be computed. In practice, not only do we consider points, but also lines and planes. The uncertainty on these primitives depends first on the manner 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 L is represented by two vectors (**l**,**m**) defined as follows. Given two points M_1 and M_2 on L, we have :

$$\mathbf{l} = \mathbf{OM_2} - \mathbf{OM_1}$$
$$\mathbf{m} = (\mathbf{OM_1} + \mathbf{OM_2})/2$$

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 3D objects.

The description built by the Stereo system consists of two parts. First, it produces the parameters of the geometric elements represented and second, the uncertainty on these parameters. This uncertainty is represented by a set of covariance matrixes.

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 in the case of planes in (Faugeras and Lustman 1986).

The resulting description of the environment, including both the geometric aspect and the 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 3-D 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 3-D) by an imperfectly known displacement **D**. We then acquire a second stereo pair, analyse it and obtain another 3-D description of another part of the environment. If the displacement **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 it should be possible to recover a better estimate of the displacement **D** and (this is done in Section 6) to construct a better estimate of the position and orientation of these primitive which have been identified in the two positions i.e. to improve the description of the environment.

We must also say something about the way we represent 3-D displacements. Every such displacement can be decomposed in an infinite number of ways as the product of a translation characterized by a vector **t** and a rotation **R** characterized by its angle θ, its origin and its axis **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 3-D displacements as the product of the group SO_3 of rotations and the group of translations of $\mathbb{R}^3$. The question of how we parametrize SO_3 is also important. For an excellent review of the different parametrizations of SO_3, see (Stuelpnagel 1964).

In our previous work (Faugeras and Hebert 1986) we have proposed to use the 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 **R**. Indeed, for every orthogonal matrix R, there exists an antisymmetric matrix **H** such that :

$$\mathbf{R} = e^{\mathbf{H}}$$

where matrix exponentials are defined as usual as $e^{\mathbf{H}} = \mathbf{I} + \mathbf{H}/1! + \mathbf{H}^2/2! + \ldots$. Matrix **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 rejected in Appendix C. There is a one to one mapping between the set of rotations $\mathbf{R}$ and the set of antisymmetric matrixes $\mathbb{H}$ for which the norm of $\mathbf{r}$ lies within $[0,2\pi]$Moreover, matrix $\mathbf{H}$ represents the cross-product with vector $\mathbf{r}$, which we denote by $c(\mathbf{r})$. By this we mean :

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

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 $\mathbf{O'M'}$ in the coordinate system associated with the second position, then :

$$\mathbf{O'M' - ROM - t = 0} \qquad (3)$$

This equation is of the form $\mathbf{f(x,a)}$, 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 rejected 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 $(M_1'M_2')$ 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 :

$$(\mathbf{Rl},\mathbf{Rm+t})$$

We can then write that:

$$\mathbf{l' \wedge Rl = 0} \qquad (4)$$

and:

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

or, using (4):

$$\mathbf{l' \wedge (m' - Rm - t) = 0} \qquad (5)$$

The geometric interpretation of equation (4) is that the transformed line is parallel to the second line. The geometric interpretation of equation (5) is that the line joining the midpoints of the transformed segment and the second segment is parallel to the second line (see figure 2).

Letting $\mathbf{x} = (\mathbf{l'}^t, \mathbf{m'}^t, \mathbf{l}^t, \mathbf{m}^t)$ and $\mathbf{a}$ as before, this is again of the form $\mathbf{f(x,a)=0}$. The equations are rejected 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 :

$$\mathbf{n' - R\,n = 0} \qquad (6)$$

and :

$$d' - d + \mathbf{t}^t \mathbf{Rn} = 0$$

or :

$$d' - d - \mathbf{t}^t \mathbf{n'} = 0 \qquad (7)$$

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

rejected 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 3D estimation of the scene has been constructed, to track the projections in one image of the geometric features used to compute the 3D displacement. This is similar in spirit to the work of Lowe who is not using our formalism (Lowe 1985). Let us derive first the case of points. We assume that we know the image coordinates u', v' of point M' such that $\mathbf{O'M'=ROM+t}$. Therefore, letting $\mathbf{x}=[\mathbf{OM}^t,u',v']^t$ and $\mathbf{a}=[\mathbf{r}^t,\mathbf{t}^t]^t$, we can use a combination of equations (2) and (3):

$$\mathbf{f(x,a)}= \begin{array}{l} (l_3(\mathbf{ROM+t}) + l_{34})u' - l_1(\mathbf{ROM+t}) - l_{14} = 0 \\ (l_3(\mathbf{ROM+t}) + l_{34})v' - l_2(\mathbf{ROM+t}) - l_{24} = 0 \end{array} \qquad (8)$$

The equations for $\partial f/\partial x$ and $\partial f/\partial a$ are rejected in Appendix G.

Let us now derive the case of lines. Let D be a 3D 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 $\mathbf{O'M_1'=ROM_1+t}$ and $\mathbf{O'M_2'=ROM_2+t}$. 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

$$\begin{vmatrix} U & U_1' & U_2' \\ V & V_1' & V_2' \\ T & T_1' & T_2' \end{vmatrix}$$

or:

$$U(V_1'T_2' - V_2'T_1')+V(T_1'U_2' - T_2'U_1')+T(U_1'V_2' - U_2'V_1')=0$$

If we match d' to a line d'' in the focal plane defined by two points p_1 and p_2 (see figure 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:

$$\mathbf{h}_1 \wedge \mathbf{h}_2 = \mathbf{0}$$

where $\mathbf{h}_1=[V_1'T_2'-V_2'T_1', T_1'U_2'-T_2'U_1', U_1'V_2'-U_2'V_1']^t$ and $\mathbf{h}_2=[v_1-v_2, u_2-u_1, u_1v_2-u_2v_1]^t$. Now, letting $\mathbf{x}=[\mathbf{OM}_1^t, \mathbf{OM}_2^t,u_1,v_1,u_2,v_2]^t$ and $\mathbf{a}$ as before, the measurement equation is:

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

the components of $\mathbf{h}_1$ can be easily computed as functions of $\mathbf{x}$ and $\mathbf{a}$. The details are rejected 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 3D description of the environment which is both in terms of its geometry (the positions and orientations of points, lines and planes) and in terms of the uncertainty of the parameters describing these primitives. Each such description is attached to a local coordinate frame.

Second, in the case 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 3D 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 Figure 4 where m is a physical point represented by the vector $\mathbf{O_iM_i}$ and the covariance matrix cov_{Mi} in frame $\mathbf{F_i}$ (i=1,...,n), and 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 $\mathbf{D_1},.....,\mathbf{D_n}$, each represented by a vector $[r_i^t\ t_i^t]^t$ and its covariance matrix $\mathbf{S_i}$ in frame i. We therefore have the following n measurement equations :

$$\mathbf{OM - R_iO_iM_i - t_i = 0}$$

which is of the form $\mathbf{f_i(x,a) = 0}$ with $\mathbf{a=[O_nM_n]}$ and $\mathbf{x_i=[r_i^t\ t_i^t\ O_iM_i^t]^t}$. 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:

$$\partial \mathbf{f_i}/\partial \mathbf{x} = [-\mathbf{K}(\mathbf{R_i},\mathbf{O_iM_i})\quad -\mathbf{I}\quad -\mathbf{R_i}]\quad \text{a 3x9 matrix}$$

and:

$$\partial \mathbf{f_i}/\partial \mathbf{a} - \mathbf{I}\ \text{a 3x3 matrix}$$

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

$$\mathbf{l \wedge R_i l_i = 0}$$

and:

$$\mathbf{l \wedge (m - R_i m_i - t_i) = 0}$$

which are of the form $\mathbf{f_i(x,a)=0}$ with $\mathbf{a=[l^t\ m^t]^t}$ and $\mathbf{x=[r_i^t\ t_i^t\ l_i^t\ m_i^t]^t}$.
The corresponding equations for the derivatives can be easily deduced from Appendixes C and E.

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.We show in Figure 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 three, three and a half, four, and four and a half meters.The stereo program described in (Ayache and Faverjon 1985) was then used to

compute the positions in 3D 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 Figure 6a projected in a vertical plane and in Figure 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 matrixes computed from the measurement equation (1) in a vertical (7a), and 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 Figures 6 and 7 indicating that the extended Kalman filtering approach is compatible with experimental evidence. Figure 8 shows the result of the estimation of the displacement between views 1 and 2 using the previous points and measurement equation (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 of 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 matrixes, the third and fourth rows show 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 matrixes 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 the points in views 1 and 2 and Figure 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.

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

Two facts can be inferred from Figures 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.

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 Figures 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 Figure 6, i.e we show the projection of the reconstructed 3D segments in a vertical plane (Figures 16a and 17a), and in a horizontal plane (Figures 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 Figure 7, i.e we show a representation of the covariance matrixes of the endpoints of the reconstructed line segments computed from the measurement equation (2) in a vertical (18a and 19a), and a horizontal plane (18b and 19b).

Figure 20 shows the results of the estimation of the displacement of the robot from view 1 to view 2 by matching 3D lines using the measurement equations (4) and (5). Figure 20a shows a view from above of the segments in view 1 which are matched in view 2, Figure 20b shows the corresponding segments in view 2, and Figure 21a shows the result of applying the estimated displacement to the segments in view 1. The result looks an awful lot (as it should) as Figure 20b. 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 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 matrixes. These covariance matrixes embed the knowledge that the motion is close to a horizontal plane. The third row show the estimate of the displacement obtained by matching manually two lines in views 1 and 2, thus yielding two measurement equations of the 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.

Figure 22a shows the results of updating the line segments in view 2 using the estimated displacement and the matched segments in view 1 and the method described in Section 6. The results have to be compared with Figure 20b. Some improvement is noted but the results are certainly not as impressing as the ones obtained with points on the grid pattern of Figure 5.
Therefore we have tried another updating method, basically heuristic, in which we update a segment by taking the mean of its endpoints and those of the corresponding segment in view 1 after appplying to them the estimated displacement. The results are shown in Figure 23 and are seen to be much better, especially in Figure 23b. At the time of this writing we believe that the reason for this discrepancy is that the representation we are using for 3D lines is not minimal (6 parameters instead of 4 which are sufficient). This is fine as long as the representation appears only in the observation $\mathbf{x}$ of the measurement equation as it does when we are estimating the displacement, but does not work anymore when it appears in the parameter $\mathbf{a}$, as it does when we are updating the model. A possible solution is therefore to use a minimal representation for the lines in 3D when we update the model. This will be the subject of a future paper.

8. Conclusions and discussion

We have proposed in this paper a number of 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 tool of the Kalman Filter (Extended Kalman Filtering). The second one is that of using a representation of 3D rigid displacements which is adapted to such a linearization and we have come up with the exponential representation of orthogonal matrixes. 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 in this paper a characterization of this uncertainty by covariance matrixes and shown that it could be traced down all the way to the sensor noise, i.e 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 unperfectly 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 cordinate frames.

A number of 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 in this article that lines were defined by two points. In practice, this is not the case since 3D lines are reconstructed by intersecting planes defined by the focal center of the cameras and 2D line segments obtained by mean-square approximation of edge points. Therefore the uncertainty on 3D lines is more complex than the uncertainty on pairs of 3D 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 so far are quite encouraging and we think that the basic approach is promising and worth more investigation.

Acknowledgements:

The authors want to acknowledge the contribution of Bernard Faverjon to an earlier version of this work (Faugeras-Ayache-Faverjon 1986) as well as stimulating discussions with Francis Lustman, Giorgio Toscani, and Pierre Tournassoud.

APPENDIX A : Kalman Filter Equations

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 $\mathbf{a}$ constant with respect to i :

$$\mathbf{a}_i = \mathbf{a}_{i-1}$$

Using the equations of the Kalman Filter, the new estimate and estimation covariance matrix are given by:

$$\hat{\mathbf{a}}_i = \hat{\mathbf{a}}_{i-1} + \mathbf{K}_i(\mathbf{y}_i - \mathbf{M}_i\hat{\mathbf{a}}_{i-1})$$

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

$$\mathbf{S}_i = (\mathbf{I} - \mathbf{K}_i\,\mathbf{M}_i)\mathbf{S}_{i-1}$$

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 $\hat{a}_n$ and the corresponding covariance matrix S_n.

APPENDIX B : Stereo Reconstruction

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

$$\partial f/\partial a = \begin{bmatrix} ul_3 - l_1 \\ \\ vl_3 - l_2 \end{bmatrix} \text{ a 2x3 matrix}$$

and:

$$\partial f/\partial x = \begin{bmatrix} l_3\mathbf{OM} + l_{34} & 0 \\ \\ 0 & l_3\mathbf{OM} + l_{34} \end{bmatrix}$$

APPENDIX C : Exponential Representation of Rotation Matrixes

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 $e^{\mathbf{H}} = \mathbf{I} + \mathbf{H}/1! + \mathbf{H}^2/2! + ...$We also know (Gantmacher 1977) that the eigenvalues a and b of $\bar{\mathbf{R}}$ and $\bar{\mathbf{H}}$ 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 matrix **H** are easily found to be equal to 0, $i(a^2+b^2+c^2)^{1/2}$ and $-i\ (a^2+b^2+c^2)^{1/2}$. Therefore we have:

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

Let us now consider an eigenvector of matrix **H** associated with eigenvalue 0. Since matrix **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 using the formula:

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

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, Vol. 1, pp.158) which states that if we compute the Lagrange-Sylvester polynomial p of the exponential function for the eigenvalues of **H**, i.e the polynomial such that :

$$p(0) = e^0 = 1$$

$$p(i\theta) = e^{i\theta}$$

$$p(-i\theta) = e^{-i\theta}$$

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 the vector **r**. The corresponding computation is a little painful, but is certainly worth doing.

We want to compute the 3x3 matrix:

$$\mathbf{K(R,OM)} = \frac{\mathbf{d(ROM)}}{\mathbf{dr}}$$

Letting $f(\theta)=\sin(\theta)/\theta$ and $g(\theta)=(1-\cos\theta)/\theta^2$, and using Rodrigues formula and the chain rule, we can write :

$$\begin{aligned}\mathbf{d(ROM)/\ dr} &= f'(\theta)/\theta[a\mathbf{HOM}\ \ b\mathbf{HOM}\ \ c\mathbf{HOM}]\\ &+g'(\theta)/\theta[a\mathbf{H^2OM}\ \ b\mathbf{H^2OM}\ \ c\mathbf{H^2OM}]\\ &+f(\theta)[\mathbf{i}\wedge\mathbf{OM}\ \ \mathbf{j}\wedge\mathbf{OM}\ \ \mathbf{k}\wedge\mathbf{OM}]\\ &+g(\theta)[\ \mathbf{i}\wedge(\mathbf{r}\wedge\mathbf{OM})+\mathbf{r}\wedge(\mathbf{i}\wedge\mathbf{OM})\ \ \mathbf{j}\wedge(\mathbf{r}\wedge\mathbf{OM})+\mathbf{r}\wedge(\mathbf{j}\wedge\mathbf{OM})\ \ \mathbf{k}\wedge(\mathbf{r}\wedge\mathbf{OM})+\mathbf{r}\wedge(\mathbf{k}\wedge\mathbf{OM})]\end{aligned}$$

Applying this matrix to a vector **v** yields:

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

Let us now decompose **v** as a component along **r** and in the plane perpendicular to **r**:

$$\mathbf{v} = \alpha\mathbf{r}+\mathbf{u} \quad \text{with} \quad \mathbf{r}.\mathbf{u}=0$$

We then have (taking into account the fact that $\|\mathbf{r}\|^2=\theta^2$):

$$\begin{aligned}\mathbf{K(R,OM)v} &= \alpha f'(\theta)\theta\mathbf{r}\wedge\mathbf{OM}+\alpha g'(\theta)\theta\mathbf{r}\wedge(\mathbf{r}\wedge\mathbf{OM})\\ &+\alpha f(\theta)(\mathbf{r}\wedge\mathbf{OM})+2\alpha g(\theta)\mathbf{r}\wedge(\mathbf{r}\wedge\mathbf{OM})+f(\theta)(\mathbf{u}\wedge\mathbf{OM})\\ &+ g(\theta)(\mathbf{u}\wedge(\mathbf{r}\wedge\mathbf{OM})+\mathbf{r}\wedge(\mathbf{u}\wedge\mathbf{OM})) =\\ \alpha[&\cos(\theta)\mathbf{r}\wedge\mathbf{OM}+\sin(\theta)/\theta\mathbf{r}\wedge(\mathbf{r}\wedge\mathbf{OM})]+f(\theta)(\mathbf{u}\wedge\mathbf{OM})+ g(\theta)(\mathbf{u}\wedge(\mathbf{r}\wedge\mathbf{OM})+\mathbf{r}\wedge(\mathbf{u}\wedge\mathbf{OM}))\end{aligned} \qquad (C.1)$$

It is interesting to note that the factor of α in the previous expresion is equal to $\mathbf{r}\wedge\mathbf{ROM}$. Indeed:

$$\mathbf{r}\wedge\mathbf{ROM} = \mathbf{H}(\mathbf{OM} + f(\theta)\mathbf{HOM} + g(\theta)\mathbf{H^2OM}) = \mathbf{HOM} + f(\theta)\mathbf{H^2OM} + g(\theta)\mathbf{H^3OM}$$

But, it is easy to show that:

$$\mathbf{H^3} = -\theta^2\mathbf{H}$$

Therefore:

$$\mathbf{r}\wedge\mathbf{ROM} = \cos(\theta)\mathbf{HOM} + \sin(\theta)/\theta\mathbf{H^2OM} = \cos(\theta)\mathbf{r}\wedge\mathbf{OM}+\sin(\theta)/\theta\mathbf{r}\wedge(\mathbf{r}\wedge\mathbf{OM})$$

So, the final result is:

$$\mathbf{K(R,OM)v} = \alpha\mathbf{r}\wedge\mathbf{ROM} + f(\theta)(\mathbf{u}\wedge\mathbf{OM})+ g(\theta)(\mathbf{u}\wedge(\mathbf{r}\wedge\mathbf{OM})+\mathbf{r}\wedge(\mathbf{u}\wedge\mathbf{OM})) =$$

$$\alpha \mathbf{r} \wedge \mathbf{ROM} + f(\theta)(\mathbf{u} \wedge \mathbf{OM}) + g(\theta)((\mathbf{OM}.\mathbf{u})\mathbf{r} + (\mathbf{OM}.\mathbf{r})\mathbf{u}) \qquad (C.2)$$

Notice that the second and third terms are zero when $\mathbf{u}=\mathbf{0}$, i.e when $\mathbf{r}$ and $\mathbf{v}$ are aligned or when the two rotations they represent have the same axis. In that case, the proof of (C.2) is much simpler since the two rotations commute, which is what we had assumed in (Faugeras 1986).
Indeed, using the chain rule, we can write :

$$(d(e^{\mathbf{H}}\mathbf{OM})/d\mathbf{r})\mathbf{v} = (de^{\mathbf{H}}/d\mathbf{H})[(\partial \mathbf{H}/\partial a)\mathbf{OM}, (\partial \mathbf{H}/\partial b)\mathbf{OM}, (\partial \mathbf{H}/\partial c)\mathbf{OM}]\mathbf{v} \qquad (C.3)$$

$de^{\mathbf{H}}/d\mathbf{H}$ is a linear application from the set of 3x3 matrixes into itself.But we also know that:

$$\partial \mathbf{H}/\partial a = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} = c(\mathbf{i})$$

similarly, $\partial \mathbf{H}/\partial b = c(\mathbf{j})$ and $\partial \mathbf{H}/\partial c = c(\mathbf{k})$. From this, it follows that :

$$[(\partial \mathbf{H}/\partial a)\mathbf{OM}, (\partial \mathbf{H}/\partial b)\mathbf{OM}, (\partial \mathbf{H}/\partial c)\mathbf{OM}] = -c(\mathbf{OM}) \qquad (C.4)$$

this implies that equation (C.3) can be rewritten as :

$$(d(e^{\mathbf{H}}\mathbf{OM})/d\mathbf{r})\mathbf{v} = -(de^{\mathbf{H}}/d\mathbf{H})\mathbf{OM} \wedge \mathbf{v} = (de^{\mathbf{H}}/d\mathbf{H})c(\mathbf{v})\mathbf{OM}$$

now, if $\mathbf{r}$ and $\mathbf{v}$ represent rotations with the same axis, then $\mathbf{H}=c(\mathbf{r})$ and $c(\mathbf{v})$ commute and we have :

$$(de^{\mathbf{H}}/d\mathbf{H})c(\mathbf{v}) = e^{\mathbf{H}}c(\mathbf{v})$$

This can be seen from the fact that if $\mathbf{A}$ and $\mathbf{B}$ are 3x3 matrixes, then $(d(\mathbf{A}^n)/d\mathbf{A})\mathbf{B}=n\mathbf{A}^{n-1}\mathbf{B}$ if and only if $\mathbf{A}$ and $\mathbf{B}$ commute.Therefore:

$$(d(e^{\mathbf{H}}\mathbf{OM})/d\mathbf{r})\mathbf{v} = \mathbf{R}\,(\mathbf{v} \wedge \mathbf{OM}) = \mathbf{Rv} \wedge \mathbf{ROM} = \mathbf{v} \wedge \mathbf{ROM}$$

which is the same as formula (C.2).

APPENDIX D : Matching 3D points

From equation (3) and 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 equation (C.2), we have, with the notations of Section 1:

$$\partial \mathbf{f}/\partial \mathbf{x} = [\mathbf{I} \;\; -\mathbf{R}] \quad \text{a 3x6 matrix}$$

and :

$$\partial \mathbf{f}/\partial \mathbf{a} = [-\mathbf{K}(\mathbf{R},\mathbf{OM}) \;\; -\mathbf{I}] \quad \text{a 3x6 matrix}$$

APPENDIX E : matching 3D lines

We use equations (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:

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

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

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

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

which completes the computation of $\partial f/\partial \mathbf{x}=[\partial f/\partial \mathbf{m}', \partial f/\partial \mathbf{l}', \partial f/\partial \mathbf{m}, \partial f/\partial \mathbf{l}]$, a 6x12 matrix. We also have:

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

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

which completes the computation of $\partial f/\partial \mathbf{a}=[\partial f/\partial \mathbf{r}, \partial f/\partial \mathbf{t}]$, a 6x6 matrix.

APPENDIX F : Matching planes

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

$$\partial f/\partial \mathbf{x} = \begin{bmatrix} \mathbf{I} & -\mathbf{R} & 0 & 0 \\ -\mathbf{t}^t & \mathbf{0}^t & -1 & 1 \end{bmatrix} \quad \text{a 4x8 matrix}$$

and:

$$\partial f/\partial \mathbf{a} = \begin{bmatrix} -\mathbf{K}(\mathbf{R},\mathbf{n}) & \mathbf{0} \\ \mathbf{0}^t & -\mathbf{n}'^t \end{bmatrix} \quad \text{a 4x6 matrix}$$

where $\mathbf{0}$ is a 3x1 vector of zeros.

APPENDIX G : Matching 3D and 2D points and lines

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

$$\partial f/\partial \mathbf{a} = \begin{bmatrix} (u'\mathbf{l}_3-\mathbf{l}_1)\mathbf{K}(\mathbf{R},\mathbf{OM}) & u'\mathbf{l}_3-\mathbf{l}_1 \\ (v'\mathbf{l}_3-\mathbf{l}_2)\mathbf{K}(\mathbf{R},\mathbf{OM}) & v'\mathbf{l}_3-\mathbf{l}_2 \end{bmatrix} \quad \text{a 2x6 matrix}$$

$$\partial f/\partial x = \begin{bmatrix} (u'l_3-l_1)R & l_3(ROM+t)+l_{34} & 0 \\ (v'l_3-l_2)R & 0 & l_3(ROM+t)+l_{34} \end{bmatrix}$$ a 2x5 matrix

Let us now tackle the case of lines. Remembering that $x=[OM_1{}^t, OM_2{}^t, u_1, v_1, u_2, v_2]^t$ and $a=[r^t, t^t]^t$, we can compute the components of $\mathbf{h}_1$:

$$h_{11}= (l_2(ROM_1+t)+l_{24})(l_3(ROM_2+t)+l_{34}) - (l_2(ROM_2+t)+l_{24})(l_3(ROM_1+t)+l_{34})$$
$$h_{12}= (l_1(ROM_2+t)+l_{14})(l_3(ROM_1+t)+l_{34}) - (l_1(ROM_1+t)+l_{14})(l_3(ROM_2+t)+l_{34})$$
$$h_{13}= (l_1(ROM_1+t)+l_{14})(l_2(ROM_2+t)+l_{24}) - (l_1(ROM_2+t)+l_{14})(l_2(ROM_1+t)+l_{24})$$

Using standard rules of differential calculus, we can write:

$$\partial f/\partial x=[-c(\mathbf{h}_2)\partial h_1/\partial OM_1 \quad -c(\mathbf{h}_2)\partial h_1/\partial OM_2 \quad c(\mathbf{h}_1)\partial h_2/\partial[u_1,v_1] \quad c(\mathbf{h}_1)\partial h_2/\partial[u_2,v_2]]^t$$

which is a 3x10 matrix, and:

$$\partial f/\partial a=[-c(\mathbf{h}_2)\partial h_1/\partial r \quad -c(\mathbf{h}_2)\partial h_1/\partial t]^t$$

which is a 3x6 matrix. The various partial derivatives can be easily computed using the expressions for $\mathbf{h}_1$ and $\mathbf{h}_2$ and equation (C.2).

REFERENCES

Ayache, N., Faugeras, O.D., Faverjon, B., and Toscani, G. 1985 (October 13-16, Bellaire, USA). Matching Depth Maps obtained by Passive Stereo. *Proceedings of the Third Workshop on Computer Vision : Representation and Control*, 197-204.

Ayache, N. and Faverjon, B., 1985 (October 13-16, Bellaire, USA). A fast Stereo Vision Matcher based on Prediction and Recursive Verification of Hypotheses. *Proceedings of the Third Workshop on Computer Vision : Representation and Control* , 27-37.

Ayache, N. and Faugeras, O.D. 1986. Hyper: A New Approach for the Recognition and Positioning of Two-Dimensional Objects. IEEE Transactions on PAMI, Vol. PAMI-8, No. 1, 44-54.

Bolle, R.M. and Cooper, D.B. 1985 (December, 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 Conference on Vision for Robots* .

Brooks, R. 1985. Aspects of Mobile Robot Visual Map Making. Robotics Research, the Second International Symposium, MIT Press, 369-376.

Crowley, J.L. 1986 (April 7-10, San Francisco, USA). Representation and Maintenance of a Composite Surface Model. *IEEE Conference on Robotics and Automation* , 1455-1462.

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

Durrant-Whyte, H.F. 1986 (April 7-10, San Francisco, USA). Consistent Integration and Propagation of Disparate Sensor Observations. *Proceedings 1986 IEEE Conference on Robotics and Automation, San Francisco*, 1464-1469.

Faugeras, O.D. and Toscani, G. 1986 (Miami Beach, Florida, USA). The calibration problem for

Stereo. *Proceedings of CVPR86*, 15-20.

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 Conference on Artificial Intelligence*, 143-147.

Faugeras, O.D. and Hebert, M. 1986. The Representation, Recognition, and Locating of 3D Shapes from Range Data. The *International Journal of Robotics Research* , Vol 5, No 3, 27-52.

Faugeras, O.D., Ayache, N., and Faverjon, B. 1986 (April 7-10, San Francisco, USA). Building Visual Maps by Combining Noisy Stereo Measurements. *Proceedings 1986 IEEE Conference on Robotics and Automation, San Francisco*, 1433-1438.

Gantmacher, F.R. 1977. Matrix Theory. Chelsea, New York.

Jazwinsky, A.M. 1970. Stochastic processes and filtering theory. Academic Press.

Laumond, J.P. and Chatila, R. 1985 (Saint Louis, Missouri). Position referencing and consistent world modeling for mobile robots. *Proceedings 1985 IEEE Conference on Robotics and Automation*, 138-145.

Lowe, D. 1985. Perceptual Organization and Visual Recognition. Kluwer Academic Publishers.

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. *Journal De Mathématiques Pures et Appliquées*, 5, 1st Series, 380-440.

Smith, R.C. and Cheeseman, P. 1986. On the Representation and Estimation of Spatial Uncertainty. to appear in the *International Journal of Robotics Research* .

Stuelpnagel, J. 1964. On the Parametrization of the Three-Dimensional Rotation Group. *SIAM Review*, Vol. 6, No. 4, 422-430.

	r_x (radian)	r_y (radian)	r_z (radian)	t_x (cm)	t_y (cm)	t_z (cm)	angle (degree)
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 1 : Motion between 2 grid patterns

	r_x (radian)	r_y (radian)	r_z (radian)	t_x (cm)	t_y (cm)	t_z (cm)	angle (degree)
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

Table 2 : Motion between 2 grid patterns

	r_x (radian)	r_y (radian)	r_z (radian)	t_x (cm)	t_y (cm)	t_z (cm)	angle (degree)
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

Table 3 : Motion computed between two stereo-views of an office

Figure 1: Geometry of a standard, two cameras, Stereo system

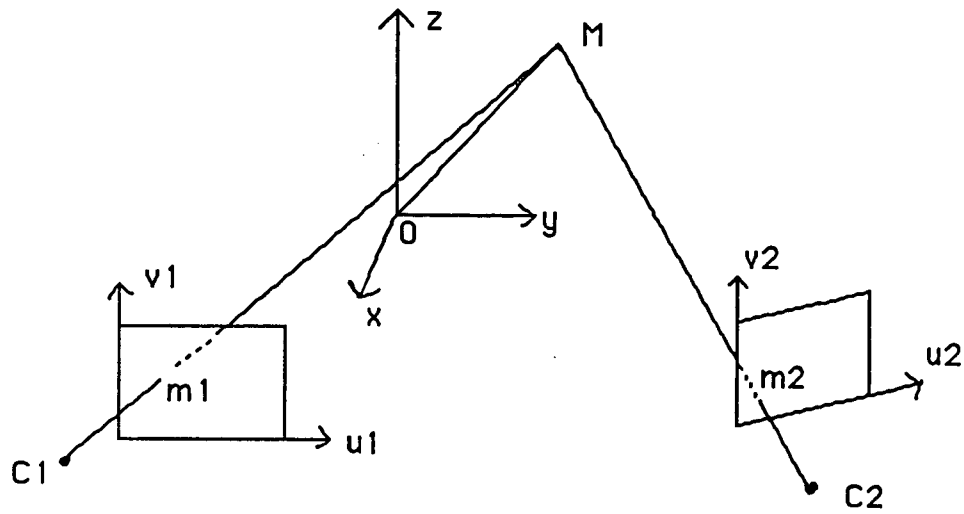


Figure 2 : D is a line reconstructed by stereo in the coordinate system (0xyz),
D' is a line reconstructed by stereo in the coordinate system (0'x'y'z')
and D* is D displaced by the rotation R and the translation $\vec{t}$.
We can write that D* is the same as D'.

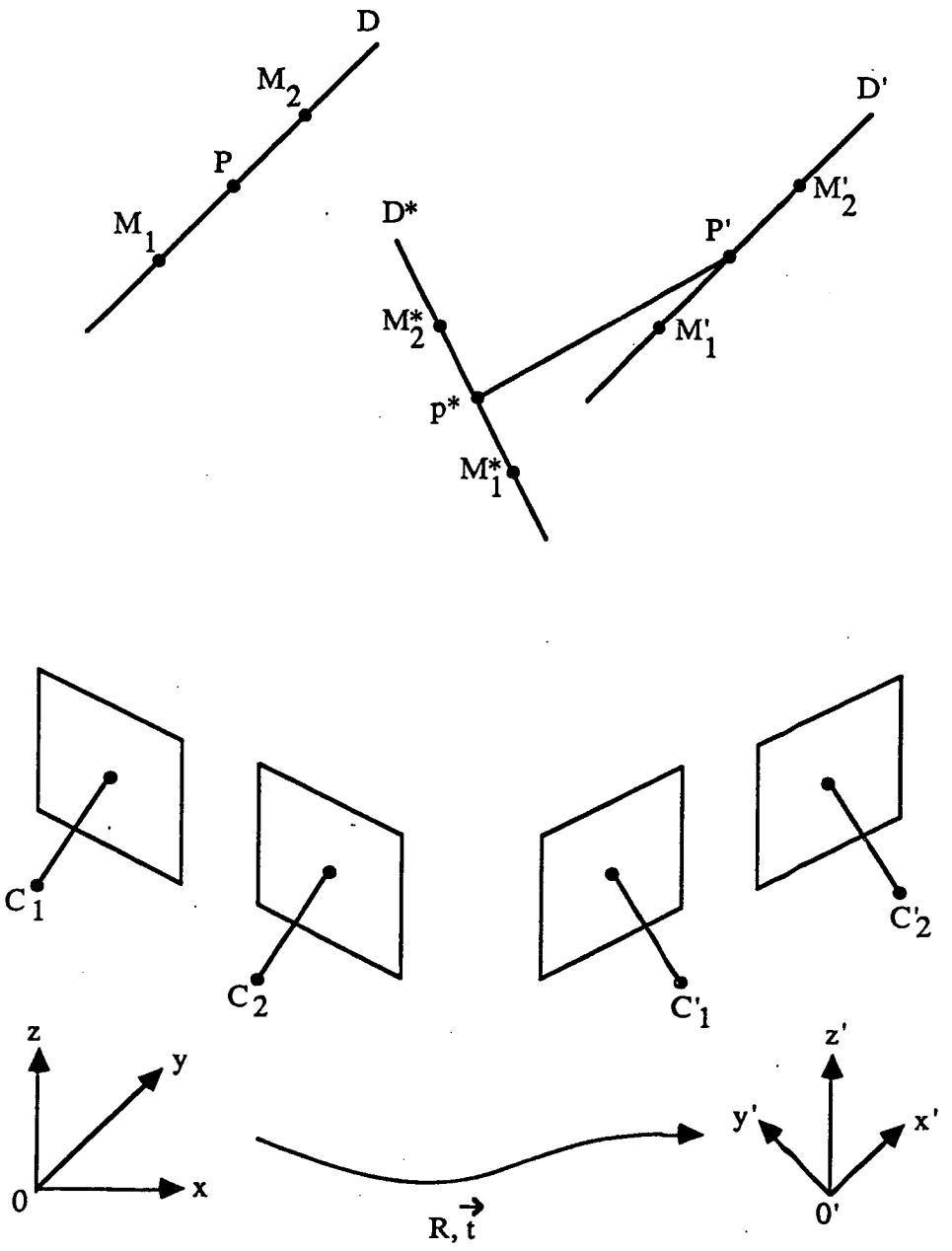
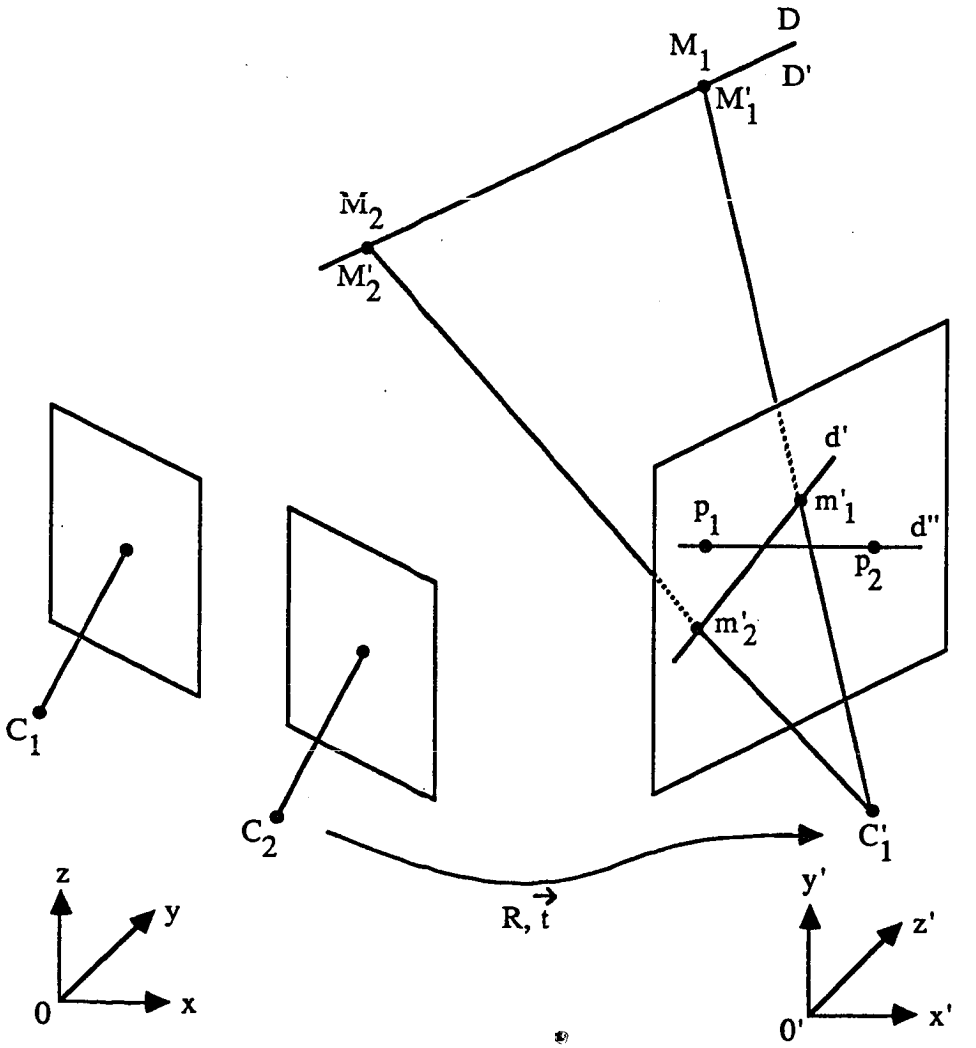


Figure 3 : After the rotation R and translation $\vec{t}$, the line D in coordinate system (0xyz) becomes line D' in coordinate system (0'x'y'z'). D' projects as d' in retina 1 and is matched to d".



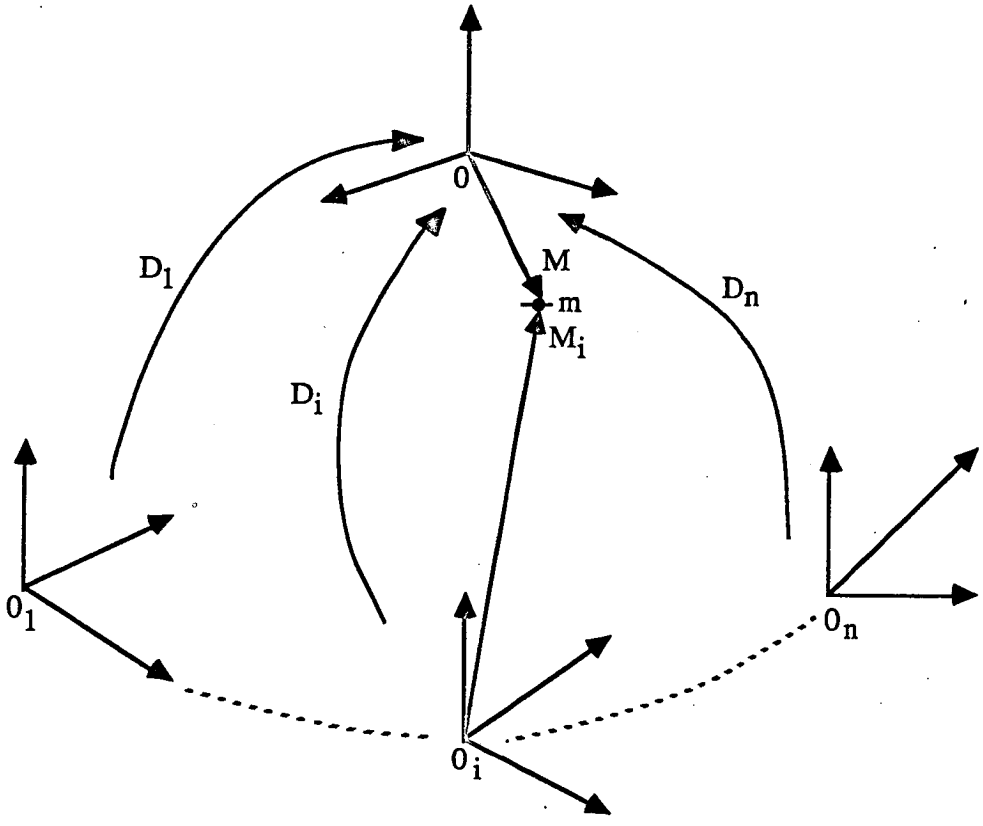


Figure 4 : The physical point m is represented by M, M_1 , ... , M_n in the various coordinate systems

Figure 5: Edges extracted from a stereo view of the grid pattern

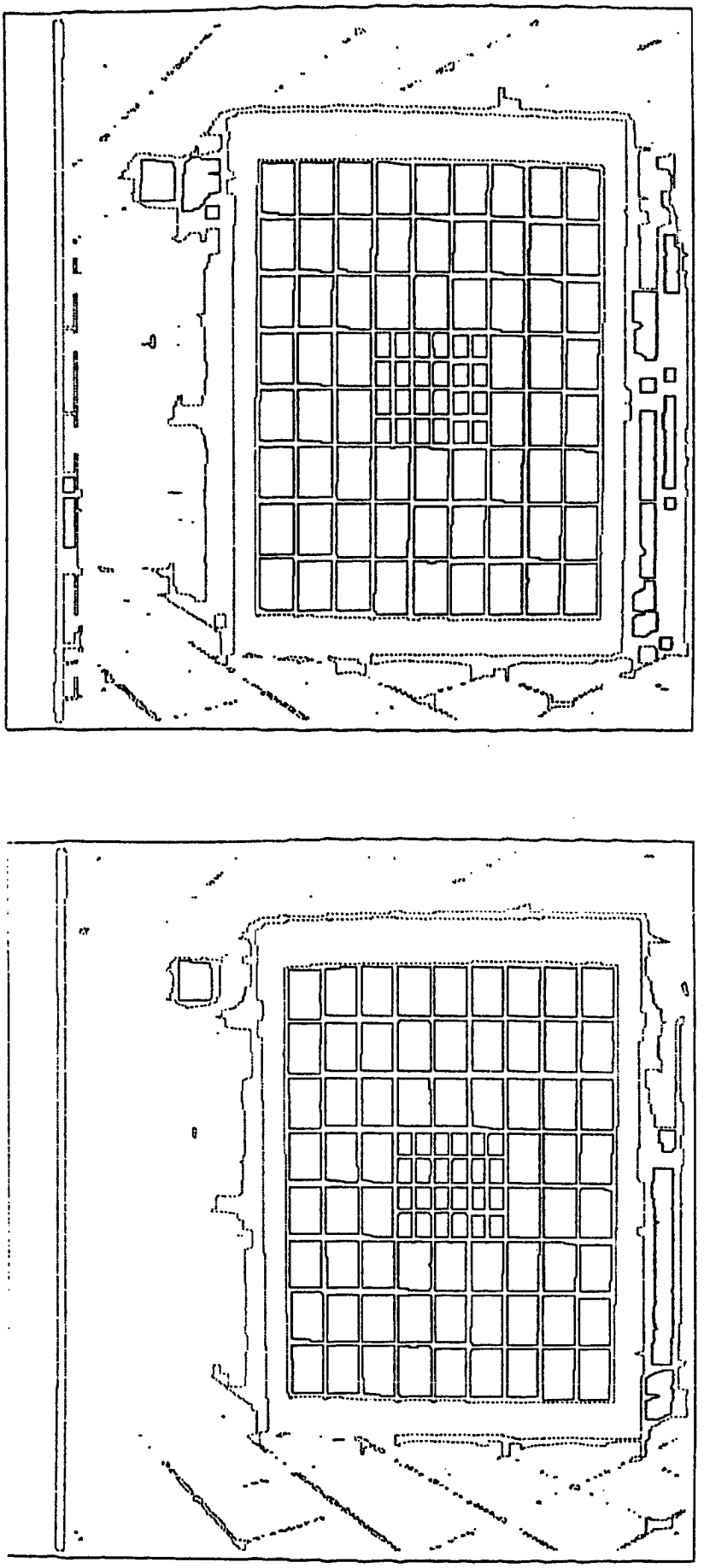
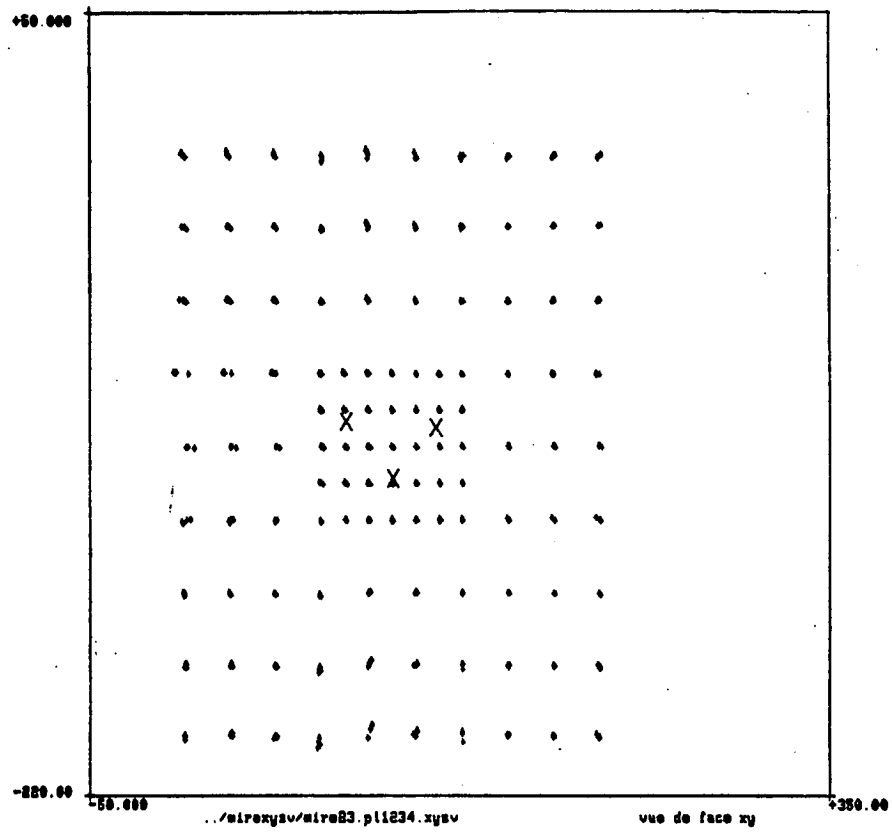


Figure 6: Horizontal and vertical projection of the reconstructed intersections of the grid pattern observed in four positions



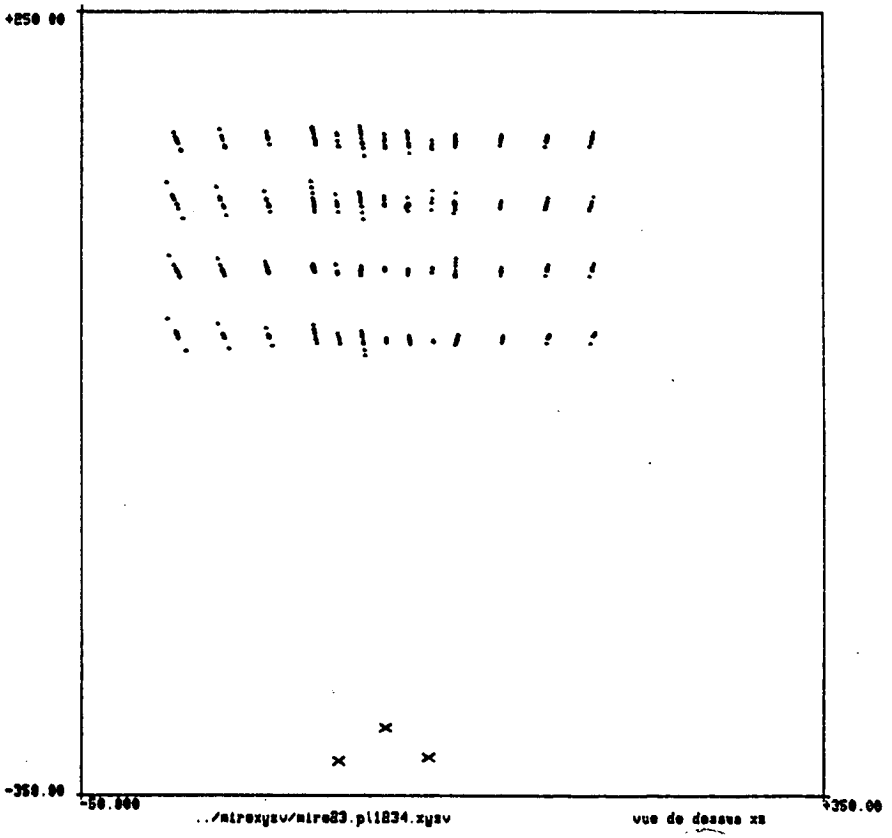
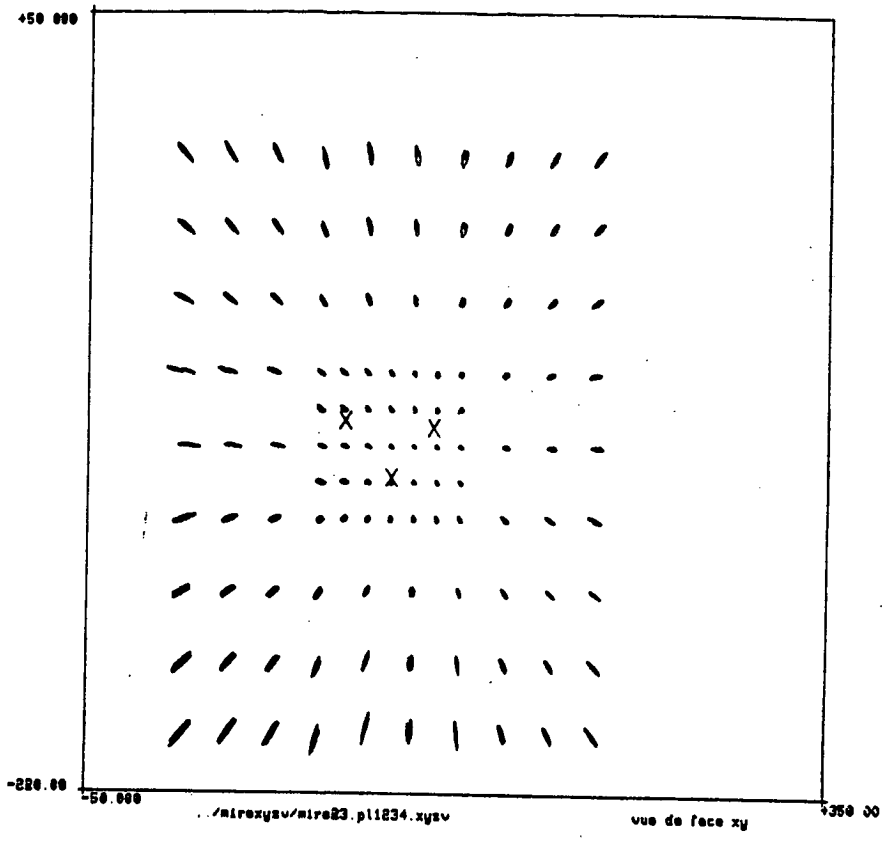


Figure 7: Covariance matrices attached to the points shown in Figure 6



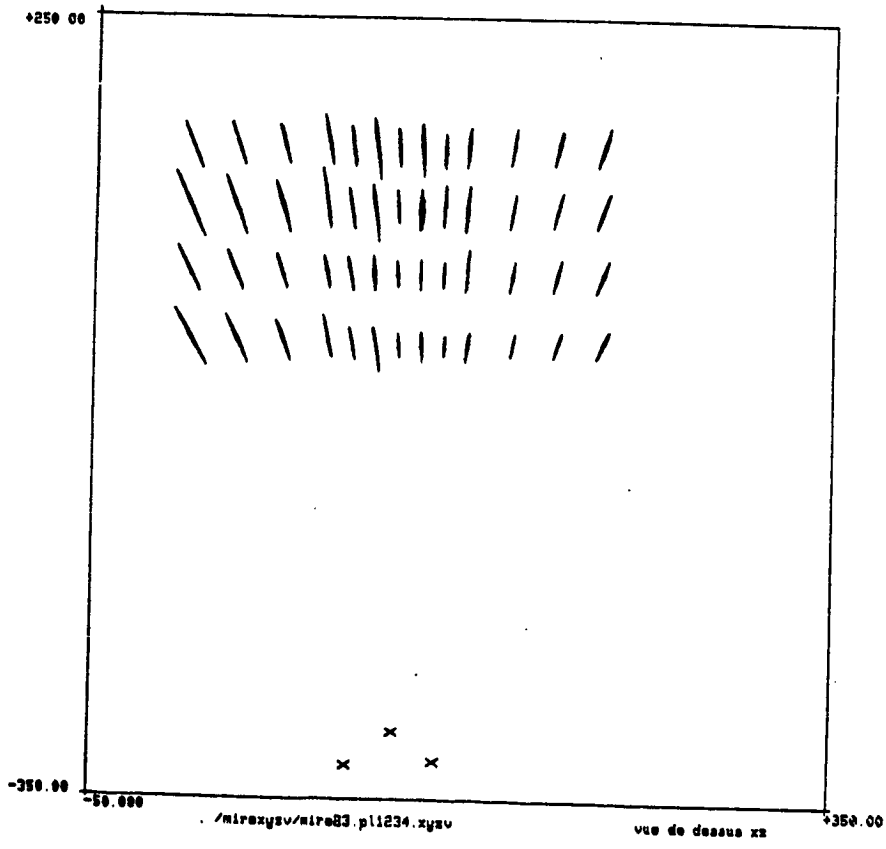
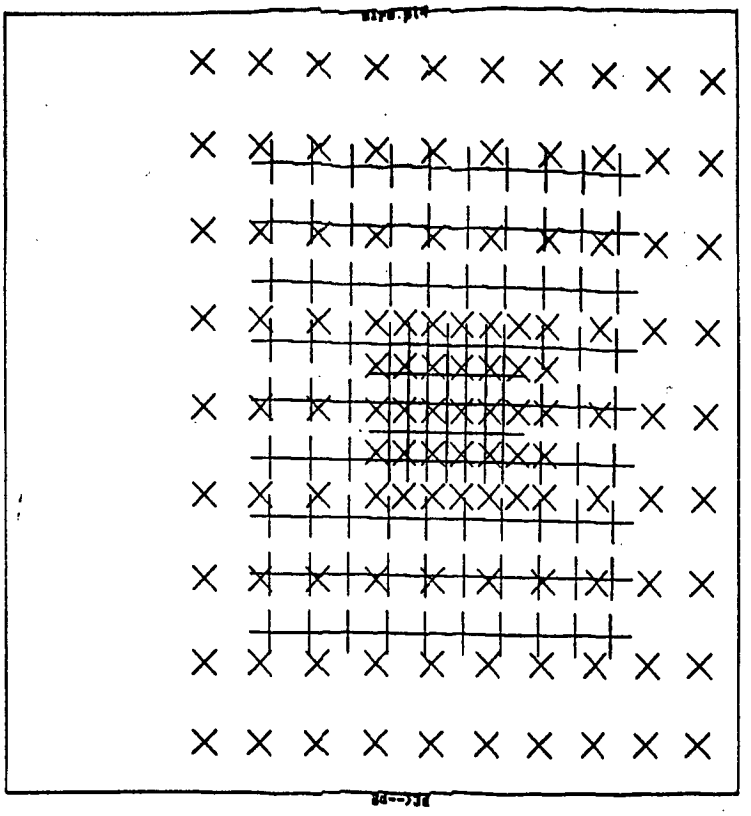


Figure 8: Images of the grid pattern observed in two different positions before and after application of the estimated displacement



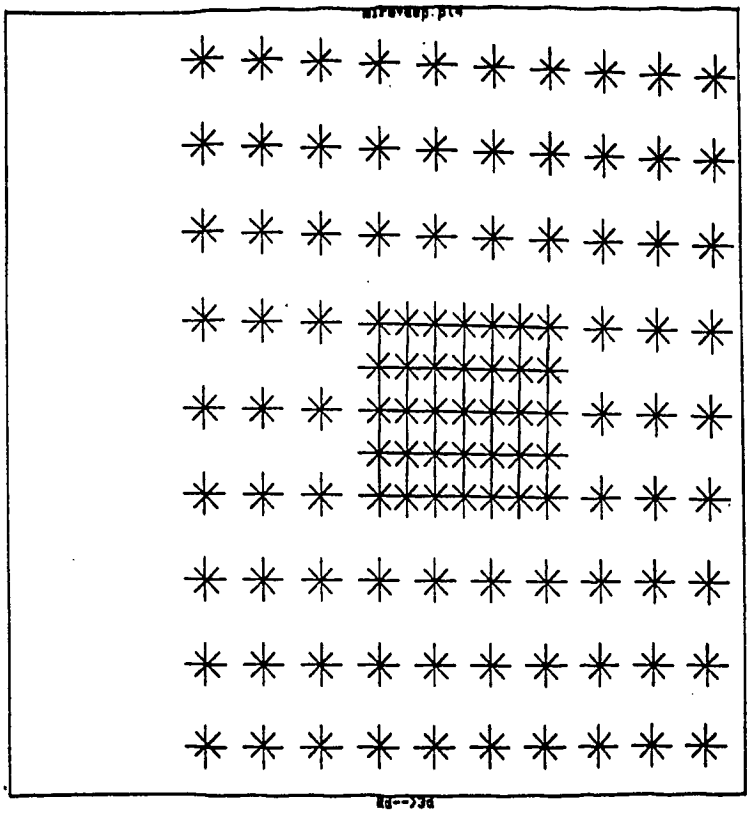
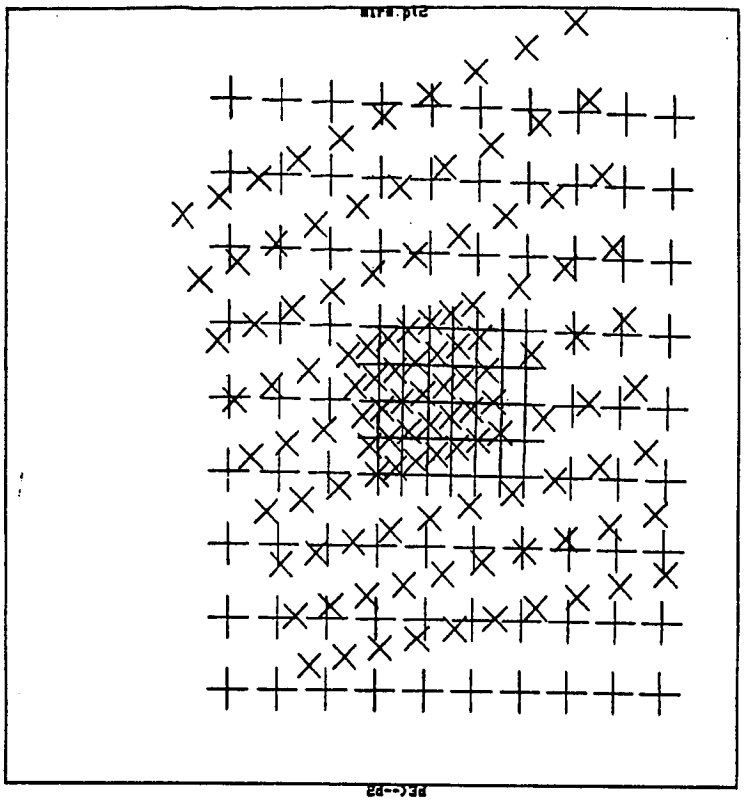


Figure 9: Images of the grid pattern observed in two different positions before and after application of another estimated displacement



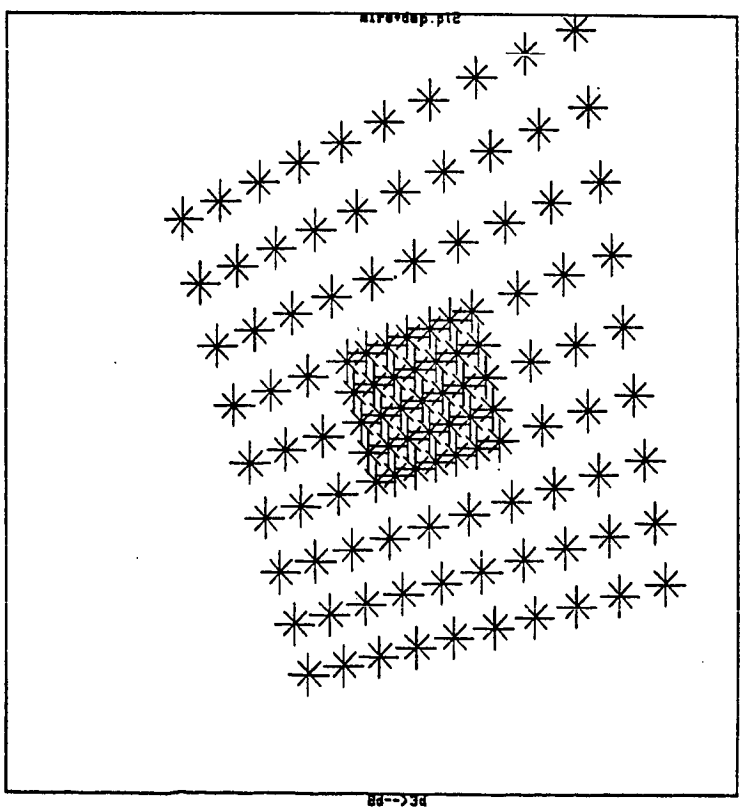
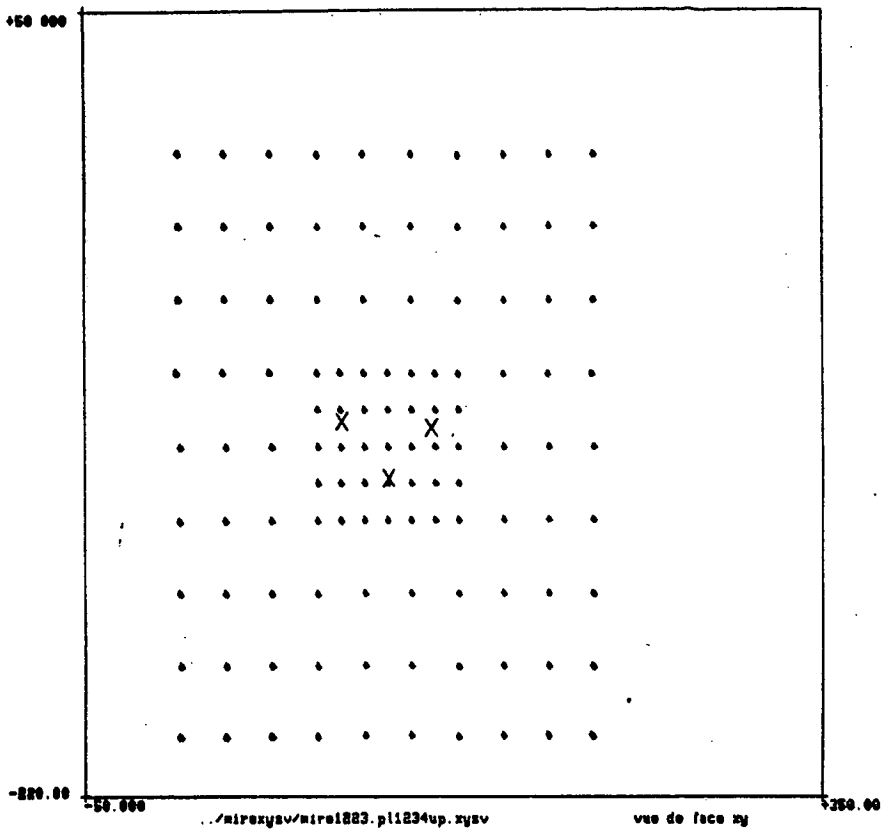


Figure 10: Fusion of the reconstructed intersections of the grid pattern



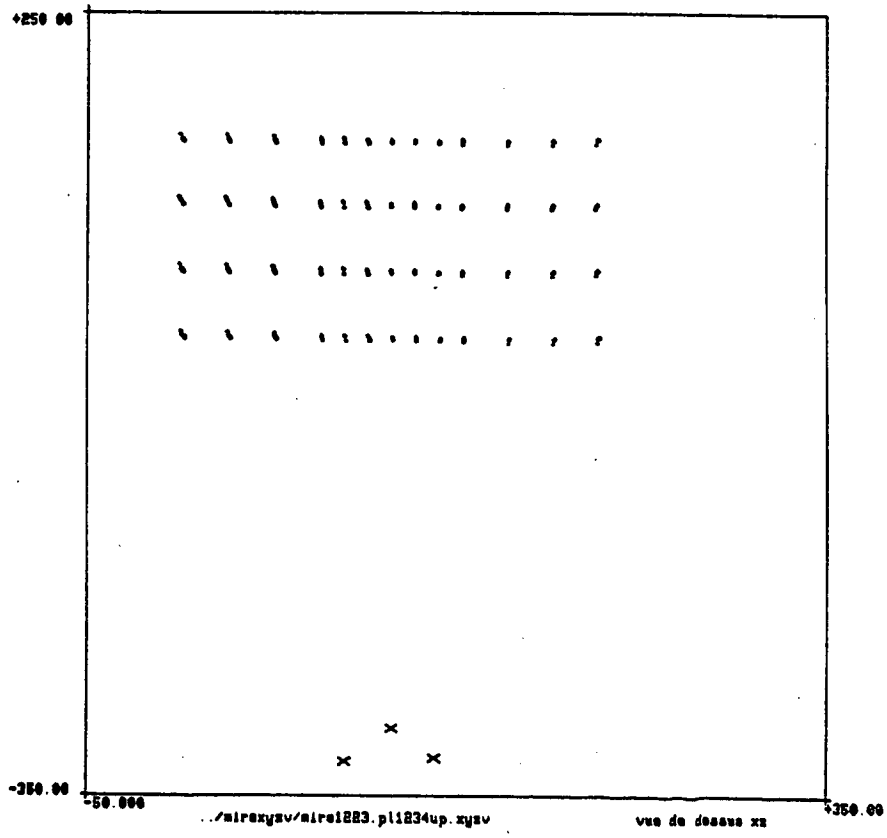
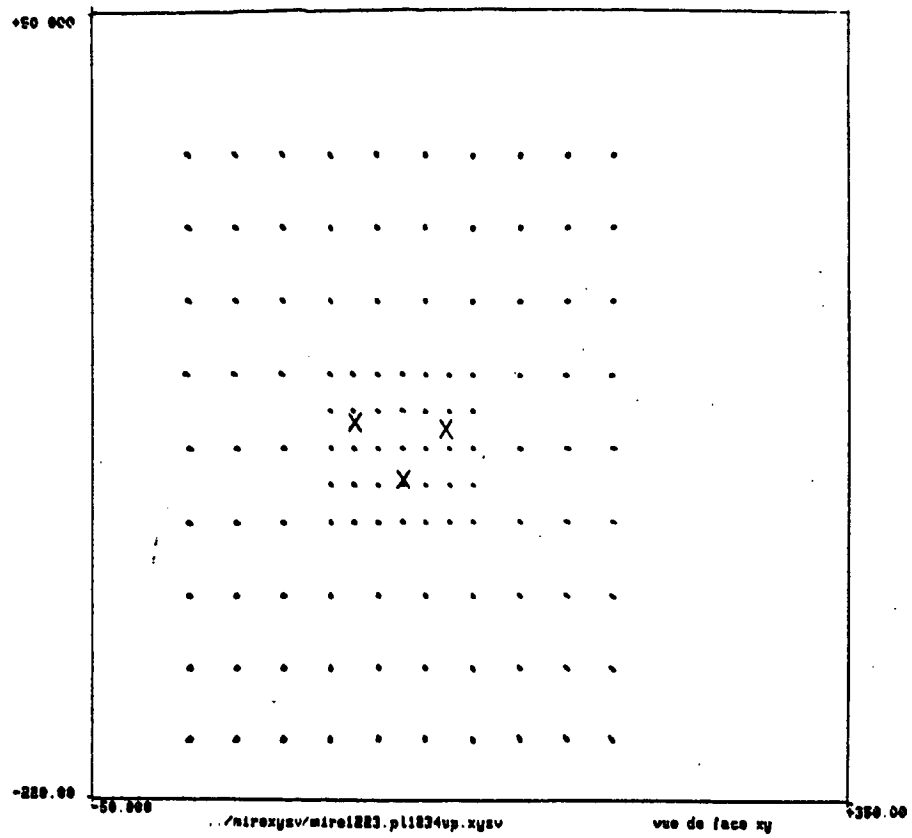


Figure 11: Covariance matrices attached to the points shown in Figure 10



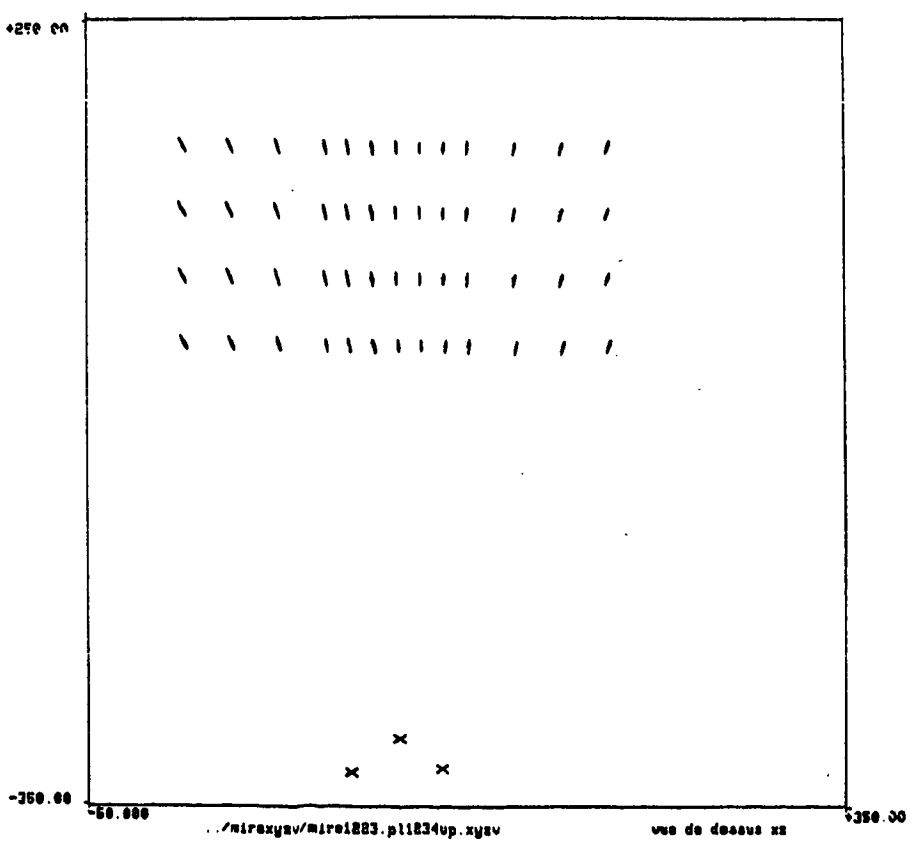


Figure 12: Polygonal approximation of the edge points of a stereo pair of an office room observed in position 1

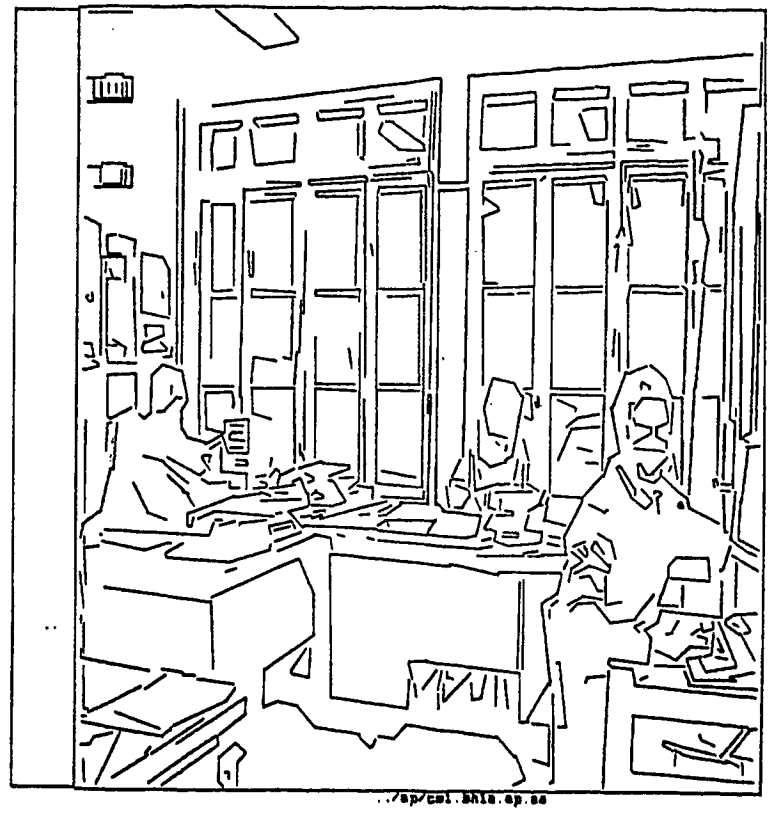
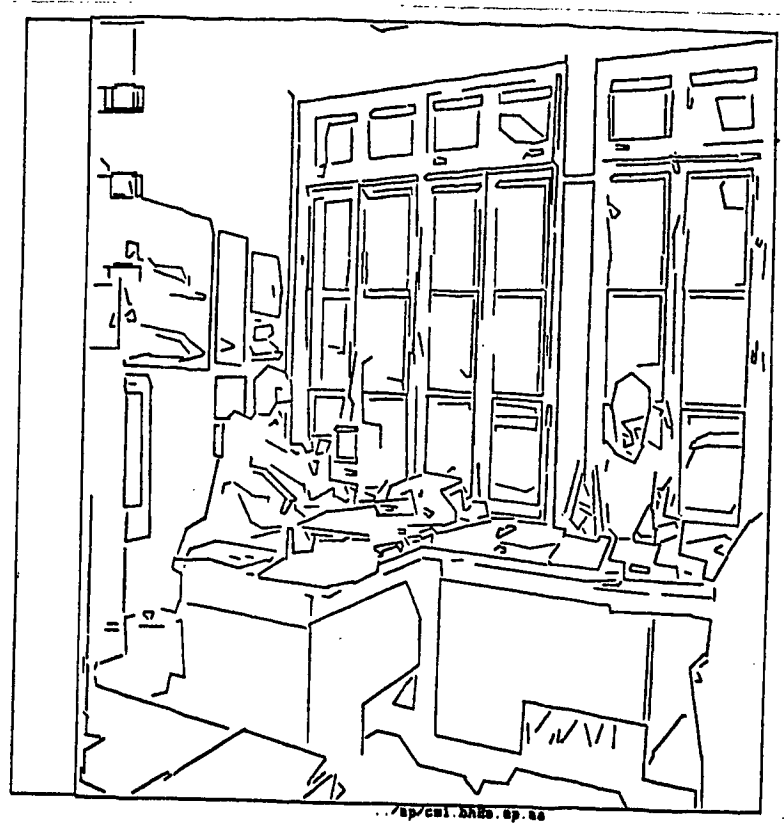


Figure 13: Polygonal approximation of the edge points of a stereo pair of the same office room observed in position 2



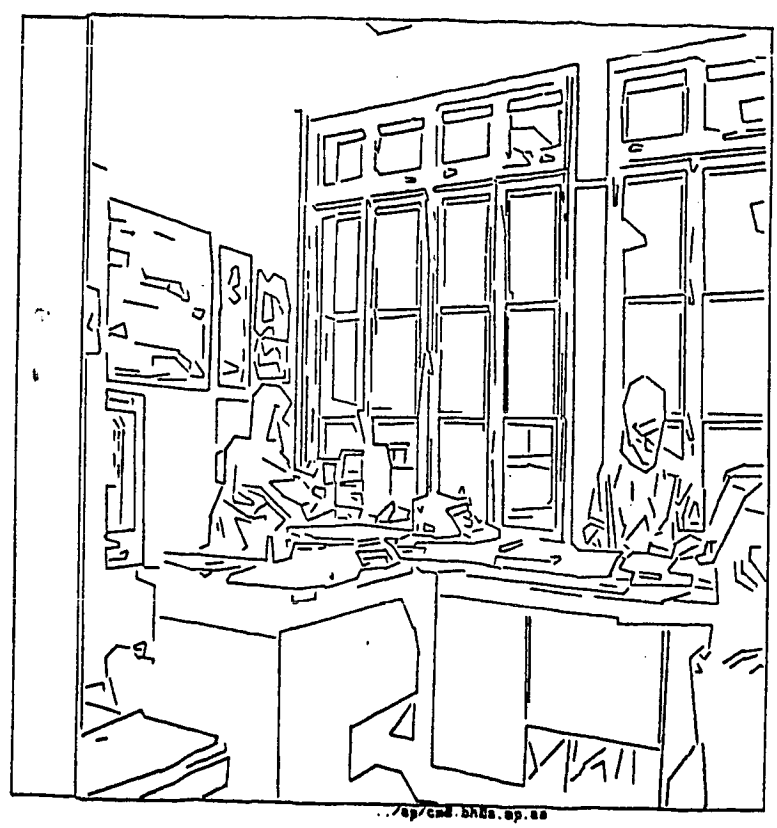
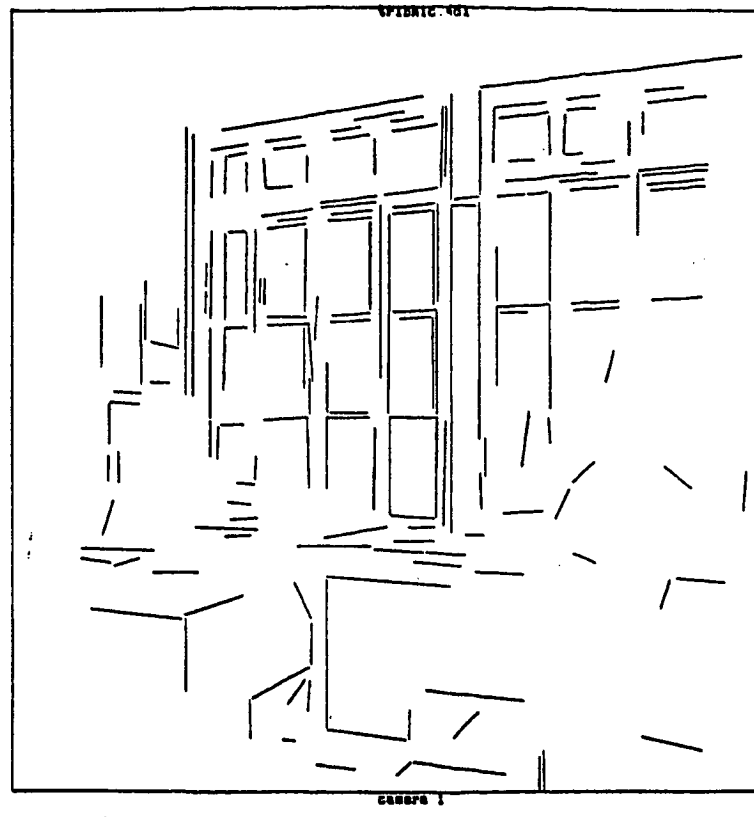


Figure 14: Edge segments matched in stereo pair of Figure 12



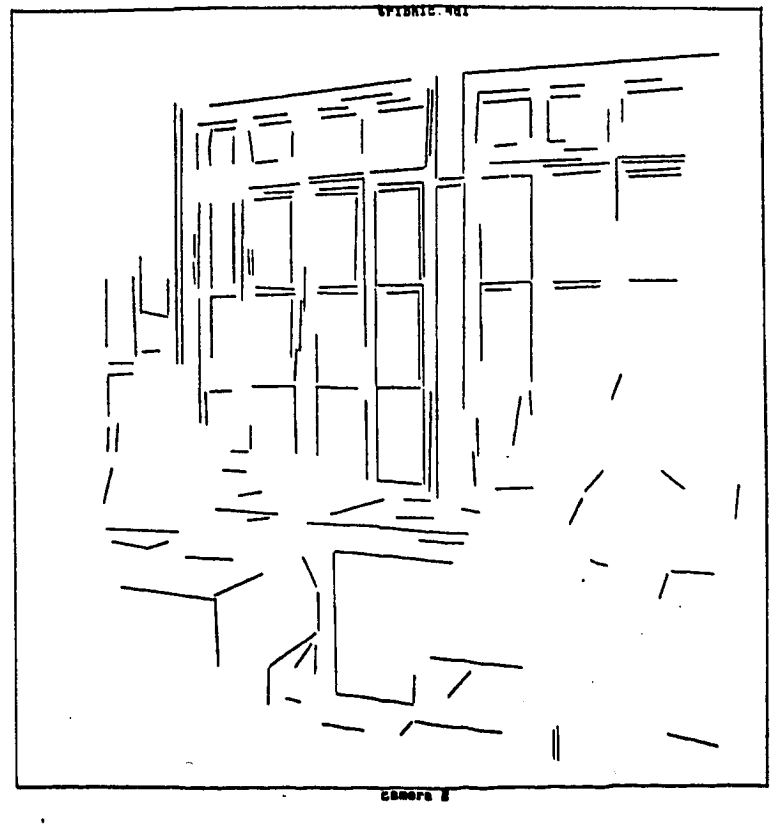
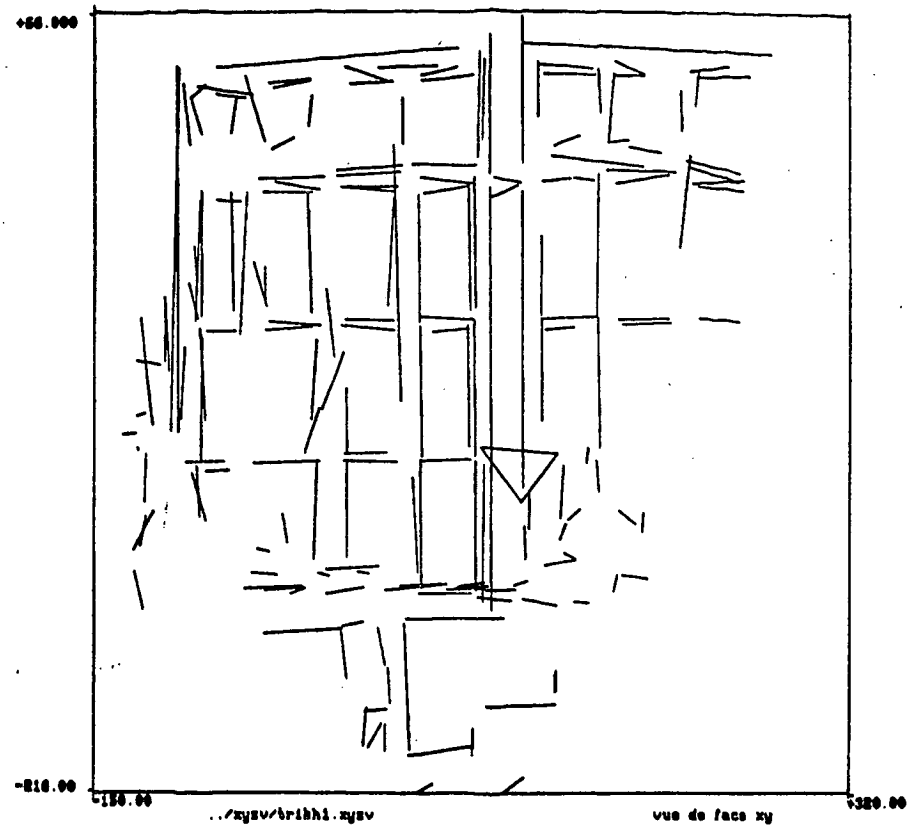


Figure 15: Edge segments matched in stereo pair of Figure 13

Figure 16: Horizontal and vertical projection of the reconstructed segments of the office room observed in position 1



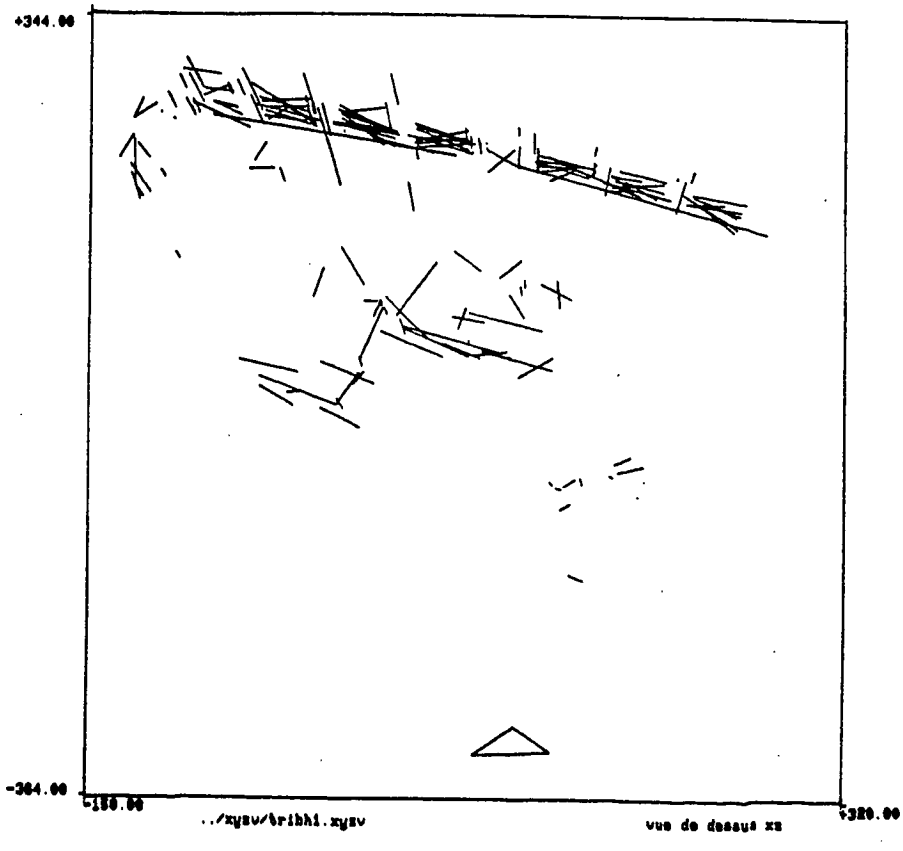
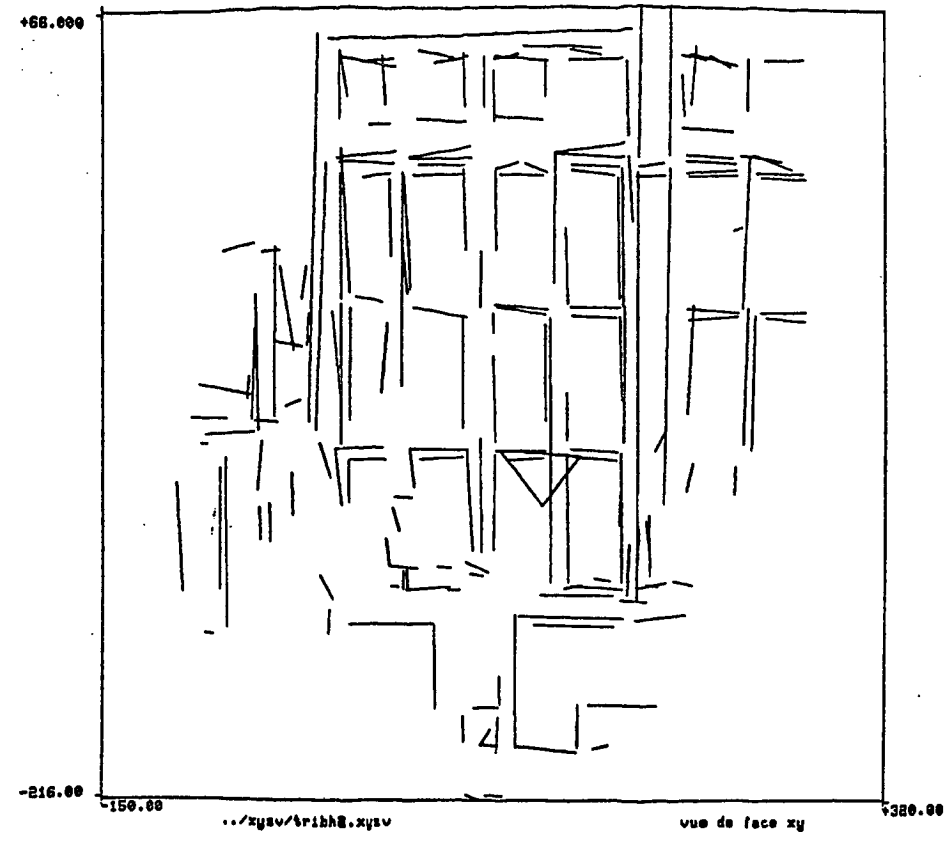


Figure 17: Horizontal and vertical projection of the reconstructed segments of the office room observed in position 2



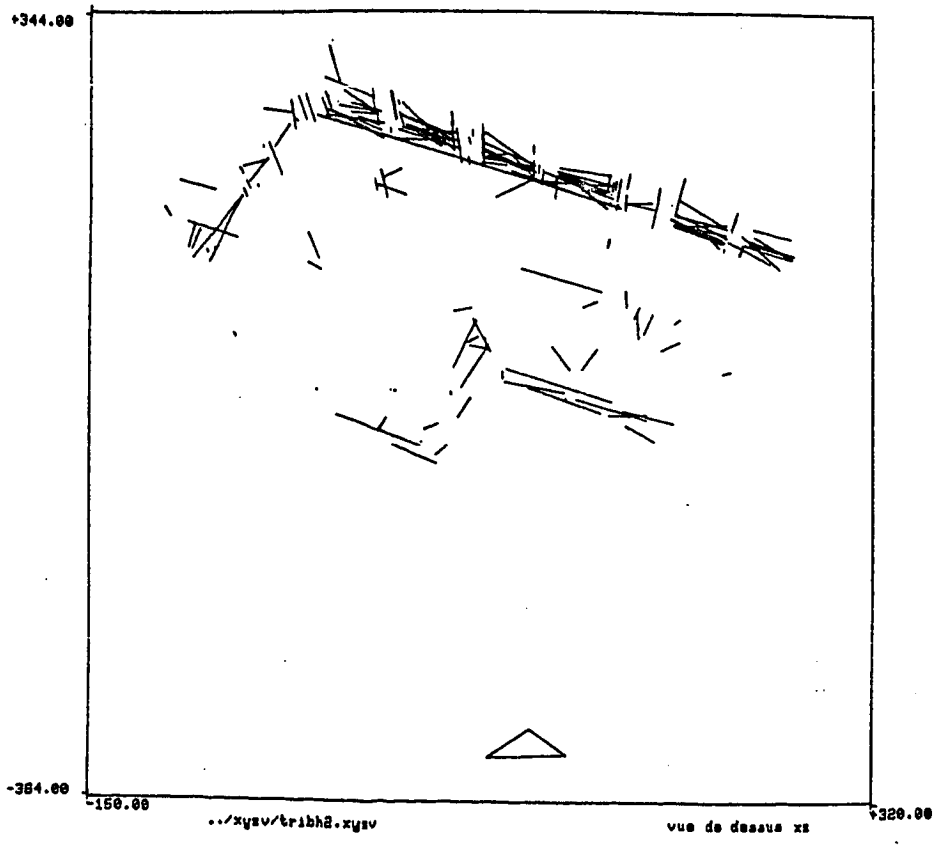
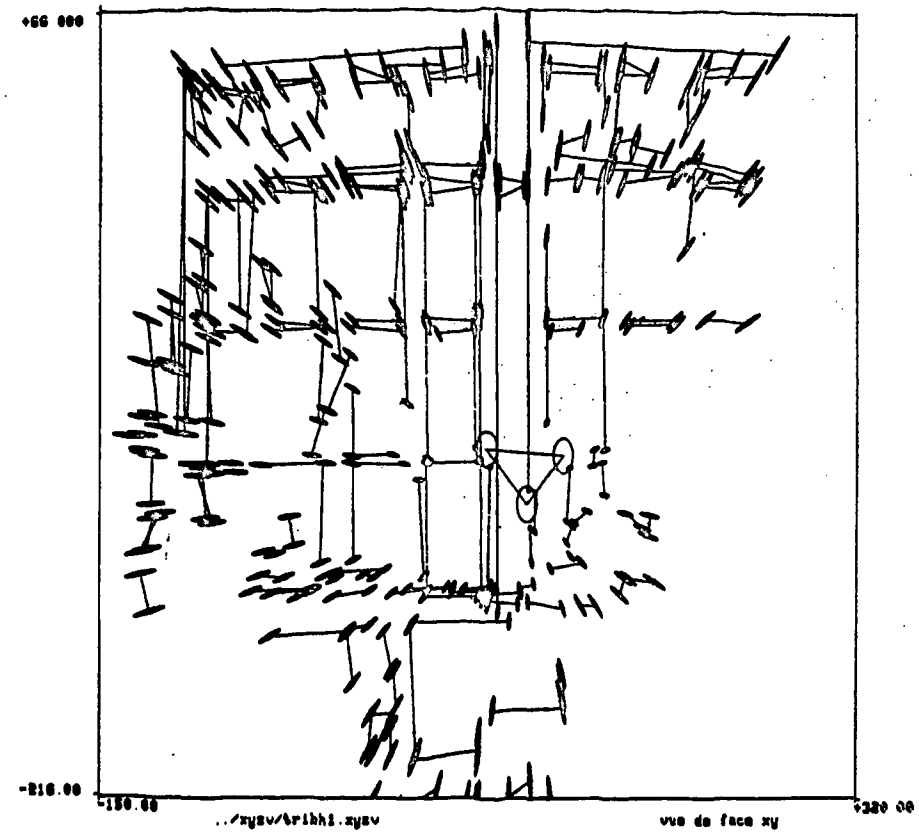


Figure 18: Covariance matrices attached to the endpoints of the reconstructed segments of the office room observed in position 1



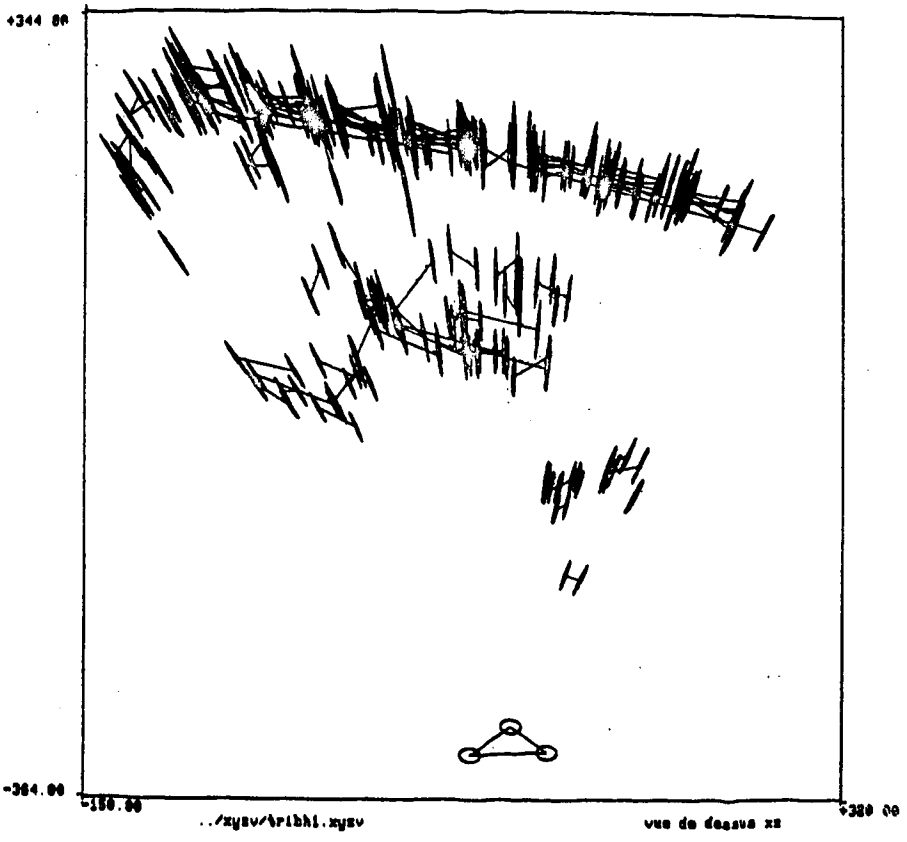
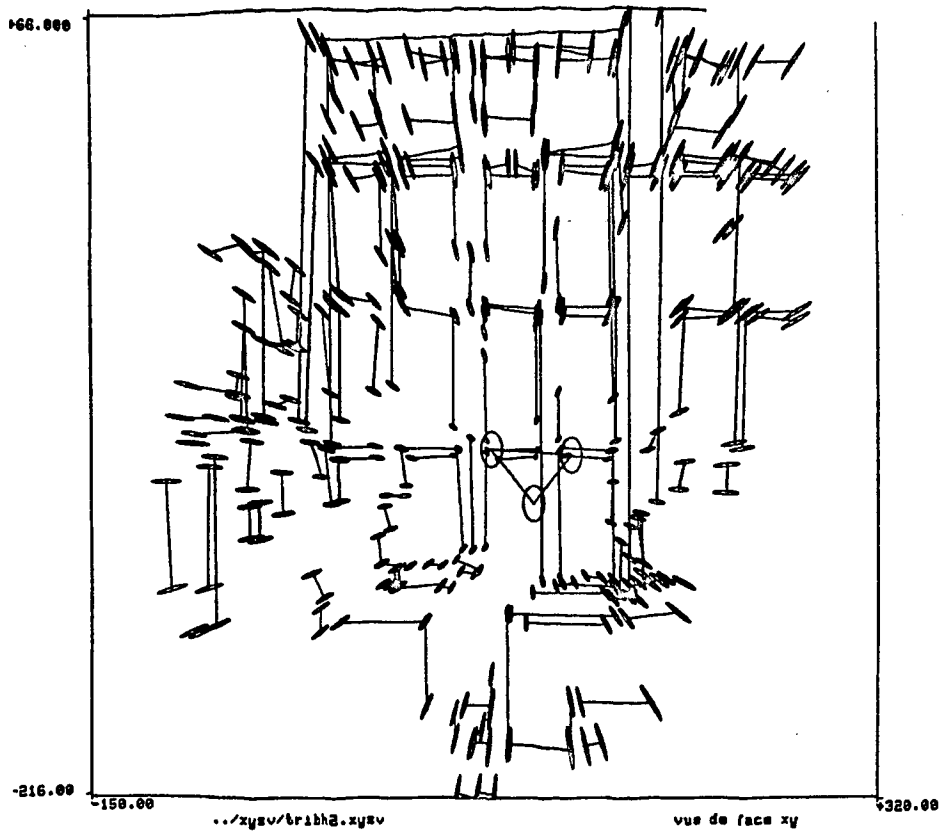


Figure 19: Covariance matrices attached to the endpoints of the reconstructed segments of the office room observed in position 2



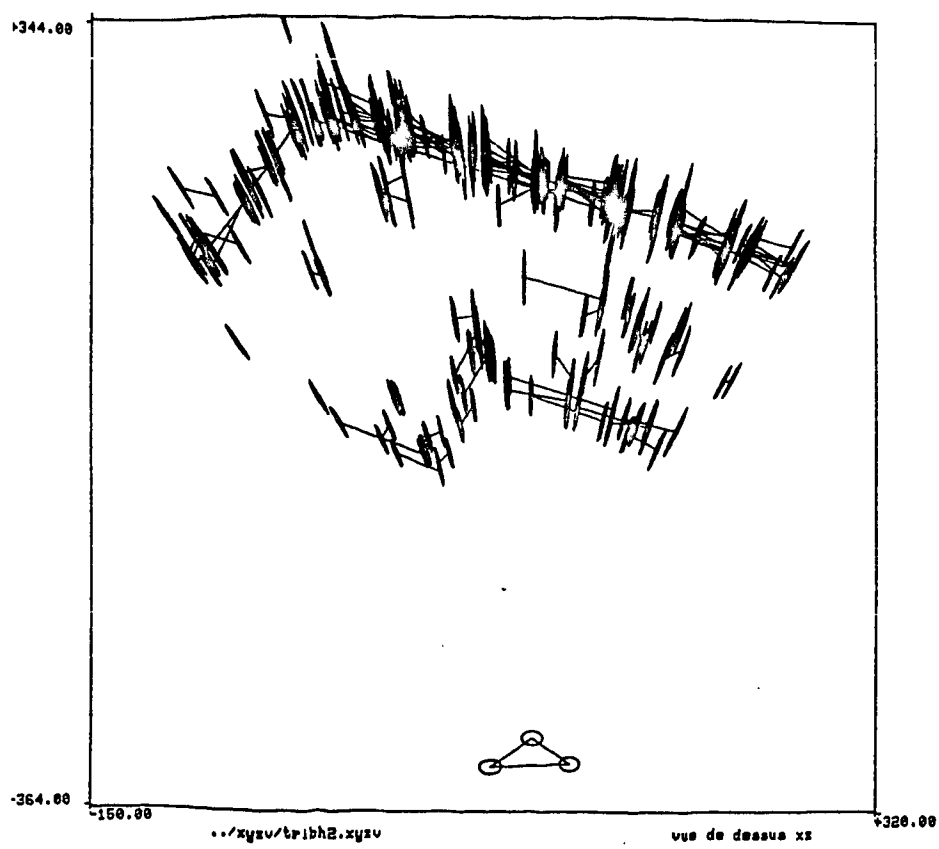
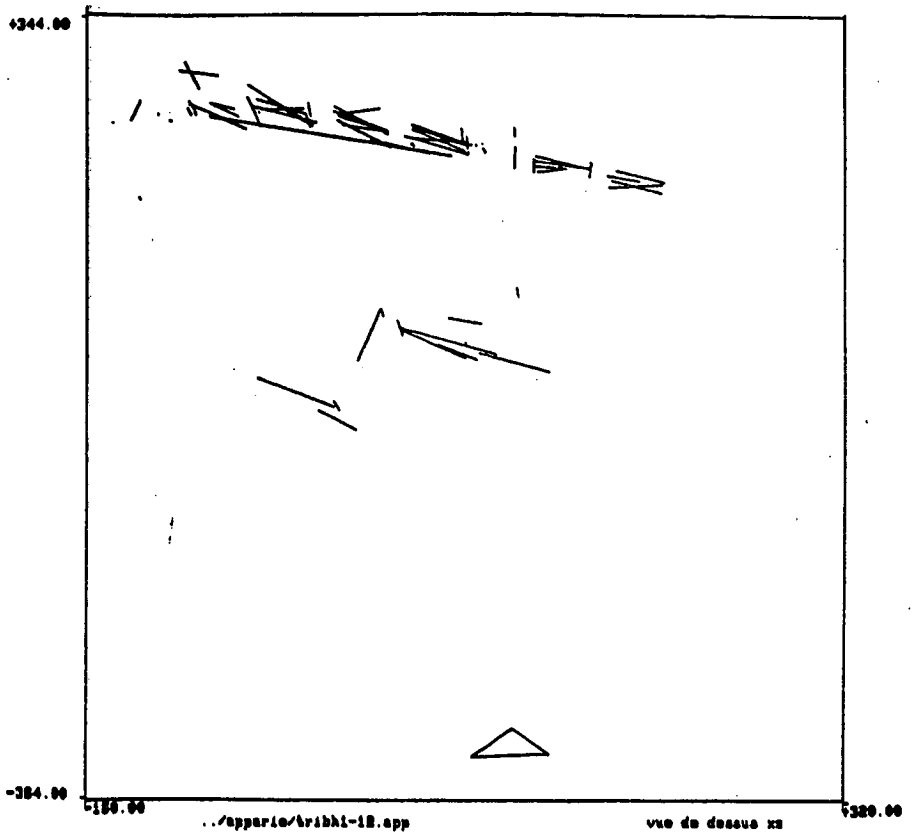


Figure 20: Reconstructed segments matched between positions 1 and 2



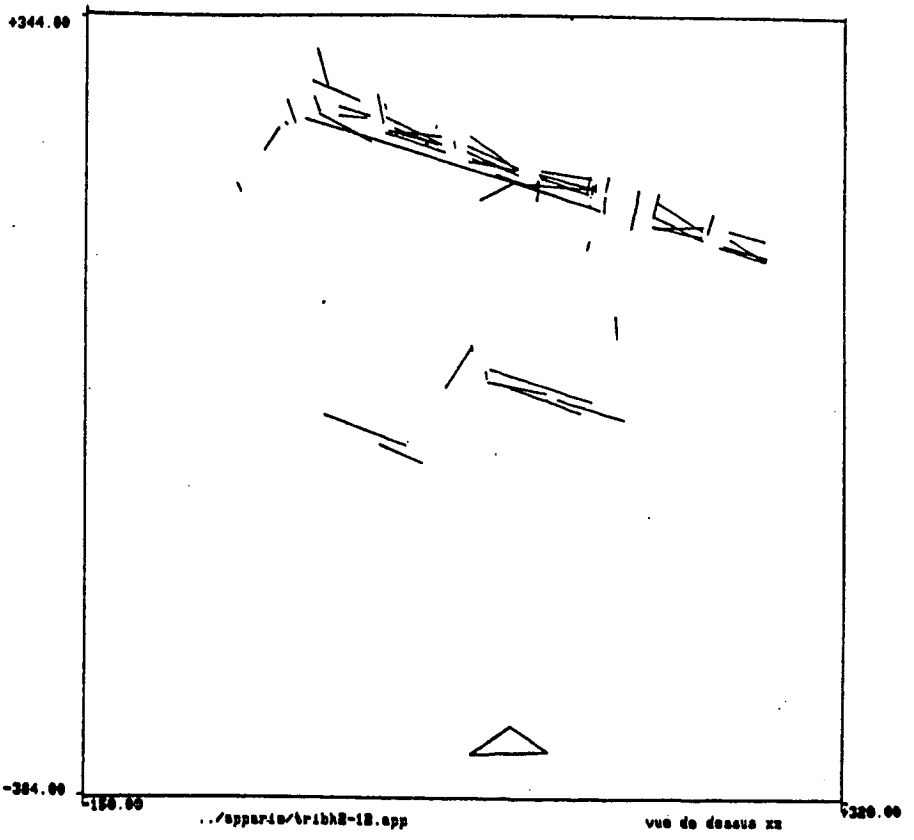
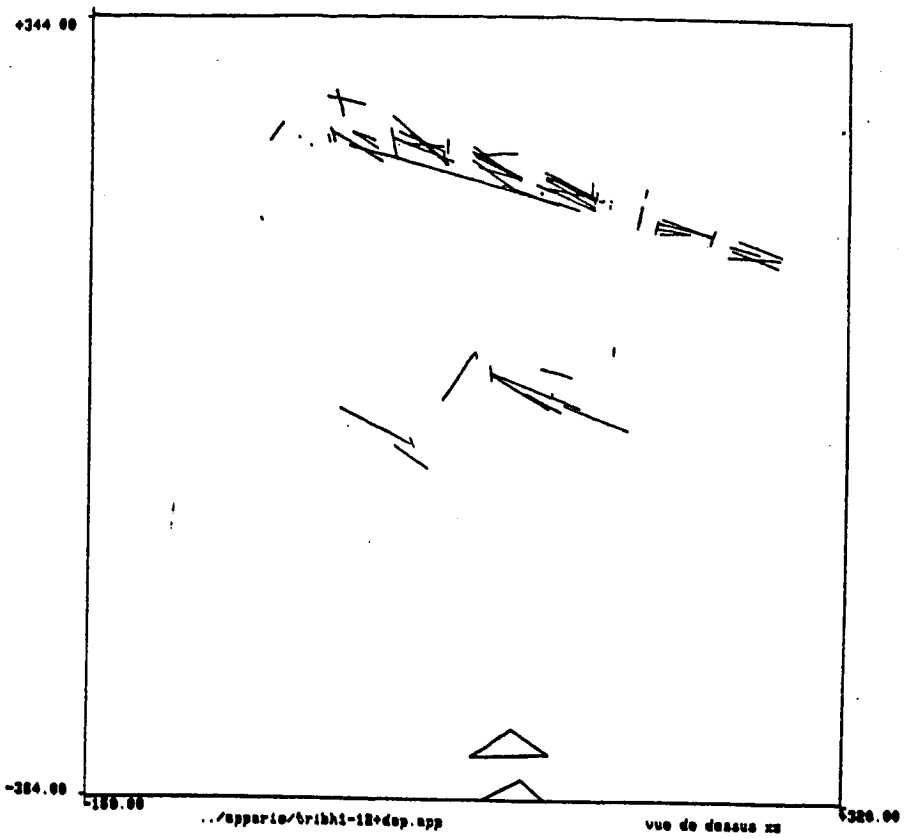


Figure 21: a) Application of the estimated motion to the segments of position 1: the triangles show the estimated motion of the robot



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).

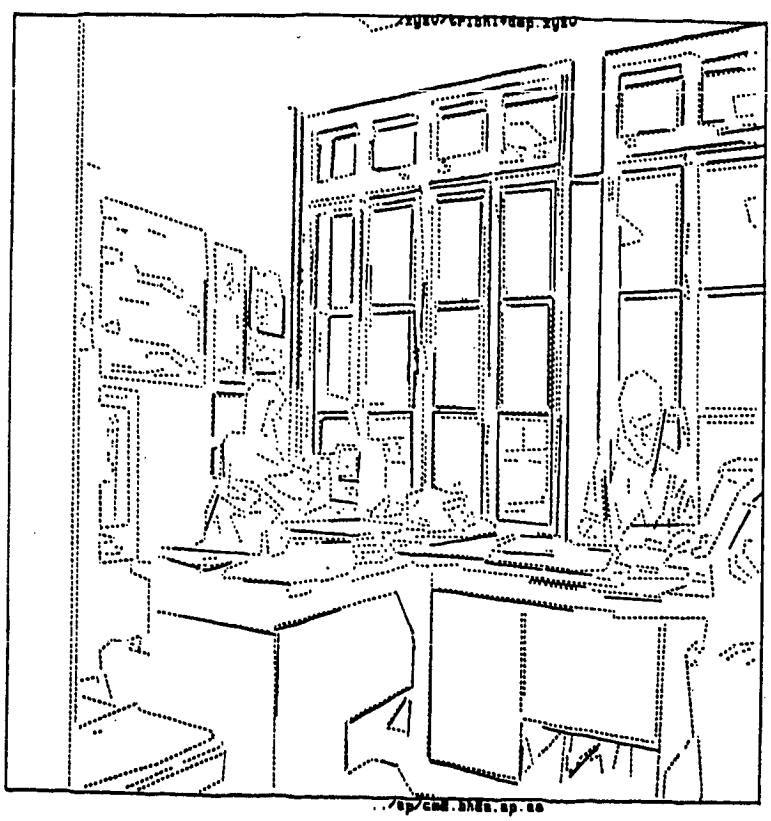
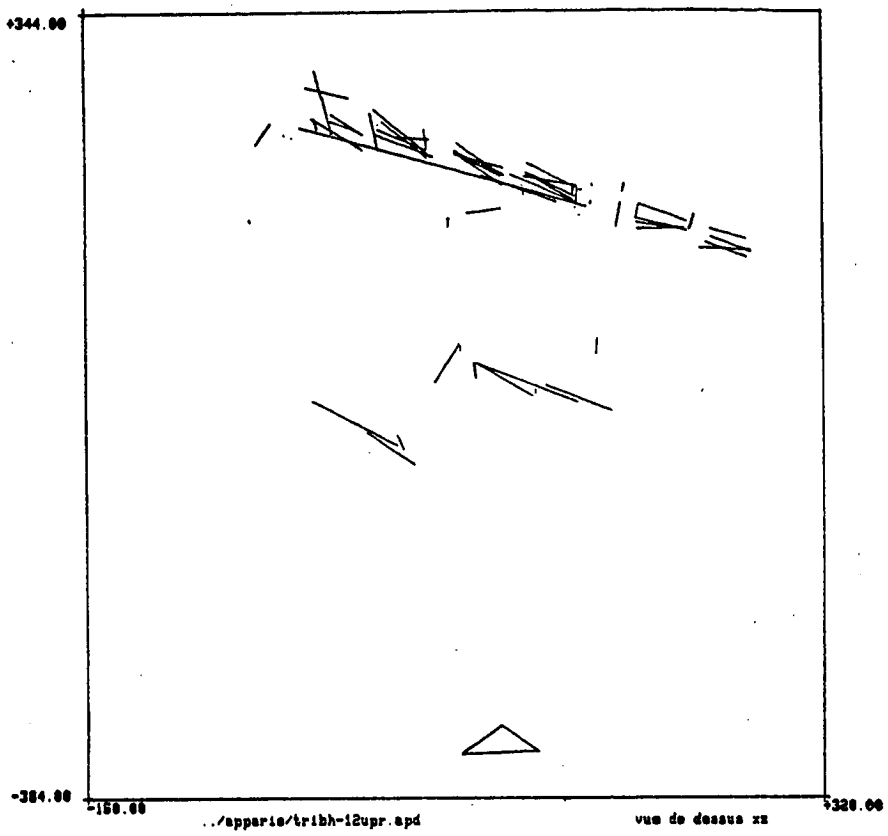


Figure 22: Fusion in position 2 of the segments reconstructed in positions 1 and 2.



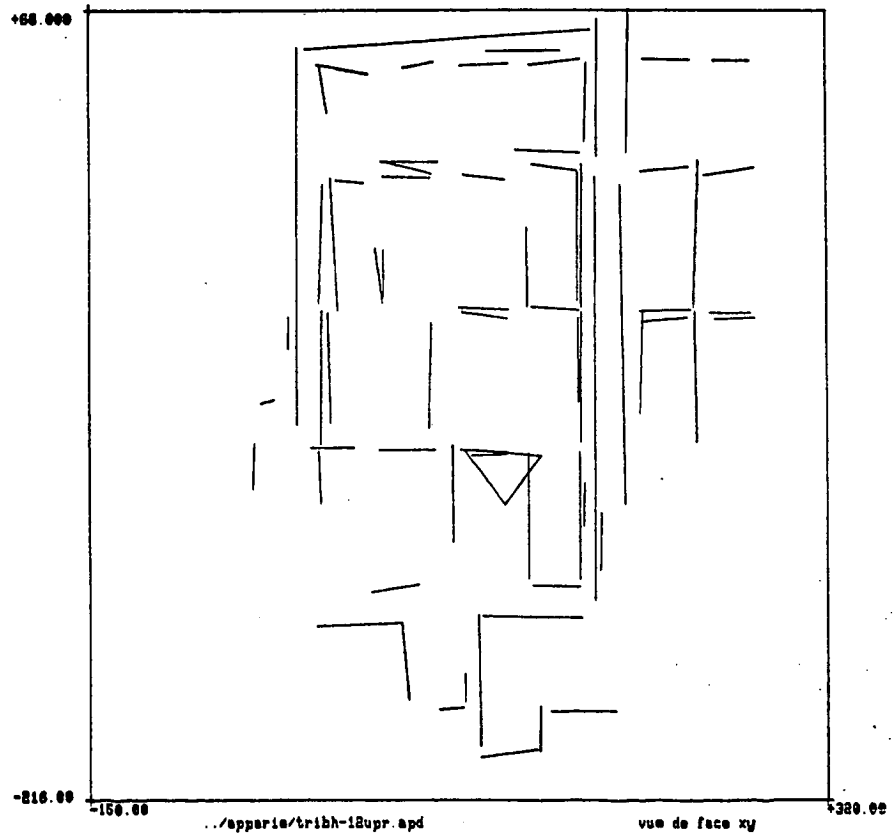
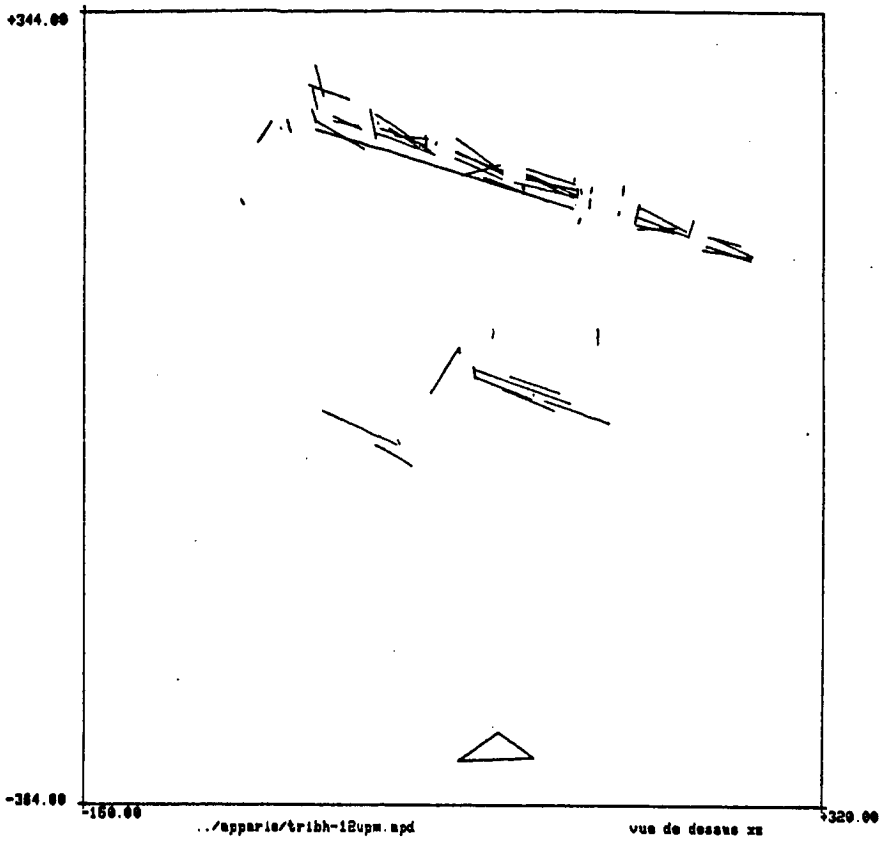
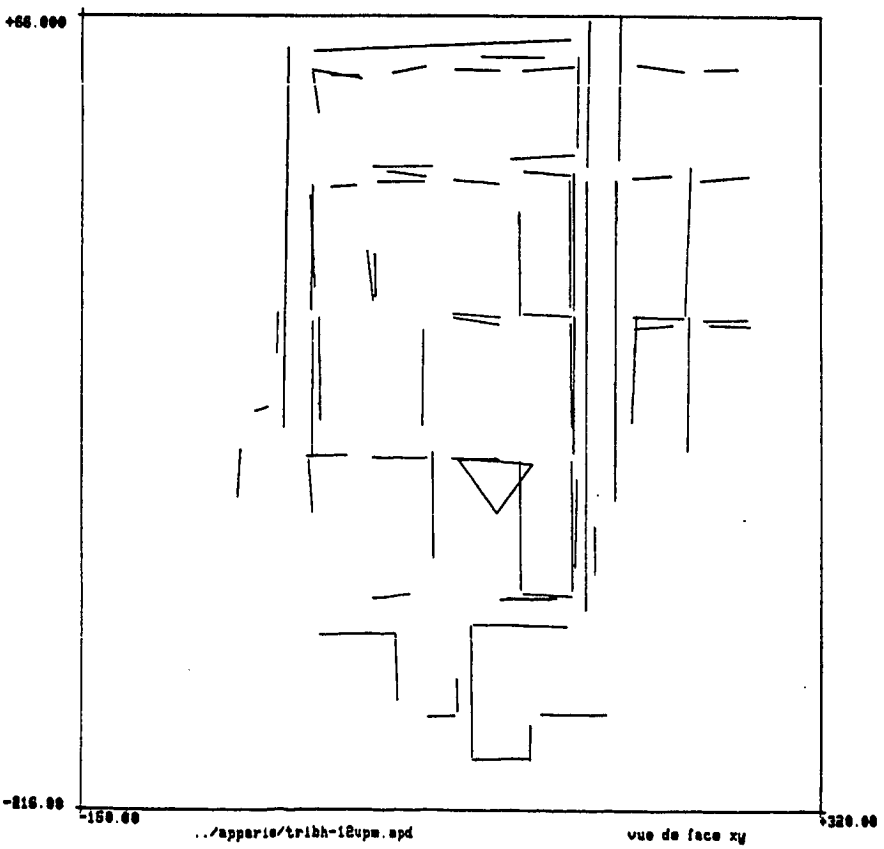


Figure 23: Fusion in position 2 of the segments reconstructed in positions 1 and 2 (another method).





Imprimé en France
par
l'Institut National de Recherche en Informatique et en Automatique