TELECOM
PARIS

COMPUTING THE PRECISION OF LOCALIZATION IN A PLANE USING PROJECTIVE INVARIANTS

TELECOM PARIS 93 D 005

1993

Patrick HORAIN

DEPARTEMENT IMAGES
Groupe Image

ECOLE NATIONALE SUPERIEURE DES TELECOMMUNICATIONS

Computing the precision of localization in a plane using projective invariants

Patrick Horain

Grup de Visió per Computador
Dept. Informàtica - Edifici C
Universitat Autònoma de Barcelona
08193 Bellaterra (Barcelona) - Spain

and

Département Images - Télécom Paris *
46, rue Barrault, 75634 Paris Cedex 13 - France

December 4, 1992

Abstract

It is well-known that projective invariants can be used to localize objects in a plan from a single perspective image. In this paper, special attention is paid to the precision of the result. A method is proposed to estimate the final uncertainty, and to minimize it. This method may be adapted to volumic invariants, and may be useful when using invariants for object recognition.

1 Introduction

Localizing objects in a 3D scene is an important topic in computer vision, because of its industrial applications.

A classical approach is stereovision. It basically consists in computing the position of scene points by triangulation. This requires the knowledge of the geometry of the adquisition system: 11 parameters for one camera, and at least the displacement to the other point of view. One of the difficulties with this approach is that camera calibration may lead to numerical instabilities [14]. Another problem is that the precision of the result strongly depends on the precision of the camera parameters [10].

In contrast, a single perspective image may be enough to get useful information on the geometry of the scene, even without knowing the imaging system parameters. For example, human vision can be quite efficient at retrieving the relative positions and sizes of objects from a single perspective photograph.

*Address for correspondance

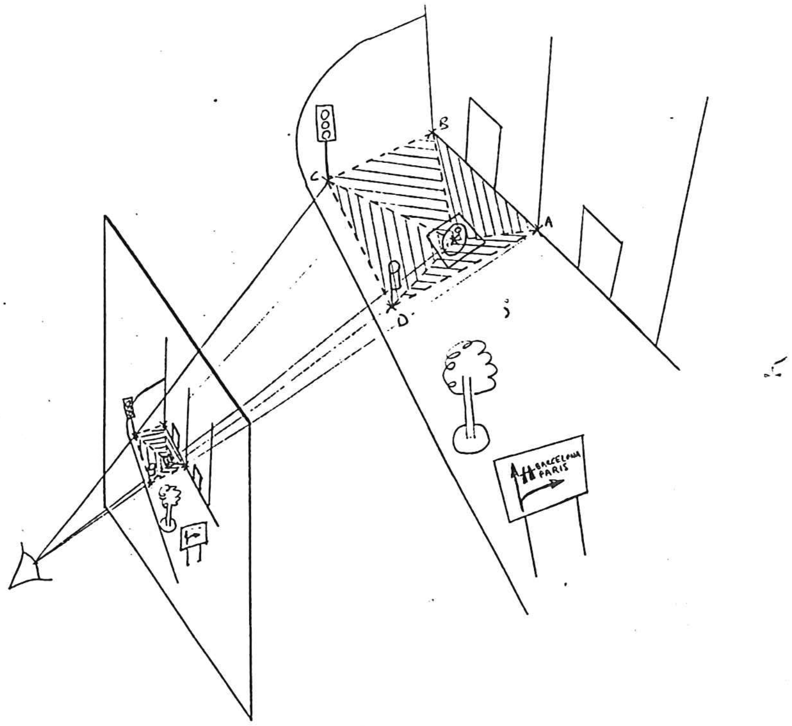

Figure 1: "Pin hole" camera model allows to compute some projective invariants, such as invariants on the areas of triangles defined between coplanar points.

Such a process is mathematically settled in the case of central projection ("pin hole" camera model), that gives various cross-ratio expressions invariant by projection: expressions based on angles in a pencil of lines, on distances between aligned points, on areas of triangles between coplanar points, or on volumes of tetrahedrons between points in the space (see for example [5, 1, 13]). These invariants can be used to derive geometrical relations between objects in the scene, from a single view and without any camera calibration.

Cross-ratios have been widely used by photo-interpreters to compute correspondances between photographs and maps [8]. Recently, computer vision scientists started to pay attention to projective invariants, and to use them for stereovision [4, 7] for object recognition [1, 2, 3, 6, 7, 11, 12].

The present work is part of a project that aims at demonstrating that computer vision can be a source of data for urban cartography [9]. The goal is to use an image to localize objects in the street (traffic signs, bins, sewer covers...) with respect to few reference points found in the database (such as the limits of buildings) (see figure 1).

Because cartographic information is essentally bidimensional, this paper will focus on the localization of objects (or their projections) on a plane (the ground). Projective invariants based on the area of triangles will be used to localize points on the ground, from a single view. A method will be proposed to minimize and estimate the uncertainty of positions, and results will be presented. This method can be extended to the utilization of volumic projective invariants, and may be useful for object recognition based on projective invariants.

2 Localization in a plane with projective invariants

Let O, A, B, C and D be five coplanar points in the scene. These points define a pencil of four lines meeting at the center point O. Provided that O, B and C on the one hand and O, A and D on the other hand are not aligned, the projective invariant of the line pencil can be defined as:

$$\mathcal{I}(O,A,B,C,D) = \frac{\sin(\widehat{AOC})\sin(\widehat{BOD})}{\sin(\widehat{BOC})\sin(\widehat{AOD})} = \frac{[\overrightarrow{OA},\overrightarrow{OC}]\,[\overrightarrow{OB},\overrightarrow{OD}]}{[\overrightarrow{OB},\overrightarrow{OC}]\,[\overrightarrow{OA},\overrightarrow{OD}]} \tag{1}$$

where the determinant $[\overrightarrow{YX},\overrightarrow{YZ}]$ is twice the signed area of the triangle (XYZ). Let o, a, b, c and d be the image points corresponding to O, A, B, C and D respectively by some (unknown) central projection. Then the following equality holds:

$$\mathcal{I}(O,A,B,C,D) = \mathcal{I}(o,a,b,c,d) \tag{2}$$

Let us suppose that the position of D in the plane of O, A, B, C and D is unknown, and that we want to obtain it from the image. Then the direction of line (OD) can be derived from equality 2:

$$\overrightarrow{OD} \;//\; \overrightarrow{OB} - k(O,A,B,C,o,a,b,c,d)\;\overrightarrow{OA} \tag{3}$$

$$\text{with } k(O,A,B,C,o,a,b,c,d) = \frac{[\overrightarrow{oa},\overrightarrow{oc}]\,[\overrightarrow{ob},\overrightarrow{od}]\,[\overrightarrow{OB},\overrightarrow{OC}]}{[\overrightarrow{ob},\overrightarrow{ob}]\,[\overrightarrow{oa},\overrightarrow{od}]\,[\overrightarrow{OA},\overrightarrow{OC}]} \tag{4}$$

Projective invariance is a property of the line pencil; so it is not possible to derive the position of D along the line (OD) from only one invariant. We need another pencil of lines meeting in some other center point O'. This pencil can be chosen either from other reference points, or from the same points used in the first cross-ratio, but taken in a different order (O' must be different from O); the image o' of O' must not be aligned with o and d.

3 Computing the precision

An important issue for cartography is the localization precision. Input data is definitely unprecise. Sub-pixel precision is hardly obtained in numerical images. The reference points in the scene may have a good precision if their topographic coordinates are read from the database; or may have poor precision if they result from some previous localization from images. Knowing the precisions of input points, what is the final precision of localization using projective invariants?

Llorens and Sanfeliu define the "sensibility" of the invariant $\mathcal{I}$ relatively to the angles between the pencil lines $\theta \in \{\widehat{AOC}, \widehat{BOD}, \widehat{BOC}, \widehat{AOD}\}$ as [5, 11]:

$$S_\theta = \frac{d\mathcal{I}/\mathcal{I}}{d\theta/\theta} = \frac{\theta}{\tan\theta} \tag{5}$$

A global sensibility for all the four angles is then derived from the partial sensibilities S_θ. Unfortunately, the invariant $\mathcal{I}$ has 2π period with respect to each angle θ, but S_θ has not; for example, $S_\theta = 1$ for $\theta = 0$, but $S_\theta \to \pm\infty$ for $\theta \to 2\pi$. More over, our input data consists of reference and image points positions; consequently, working with angles requires inverse trigonometric computing; this gives poor insight to the influence of the error on each reference point.

Morin and Mohr have used a differential method to estimate the uncertainty on the position of a point computed by stereovision, with invariants on distances along a reference line [7]. Their differential approach will be applied to the surfacic invariant of definition 1, in order to estimate the uncertainty on the computed positions.

With the assumption that errors are small in comparaison to the distance between the points, the influence of error on each reference point can be computed as the scalar products of the errors and the partial gradients (first order Taylor development of expression 4):

$$dk = \sum_{X \in \{O,A,B,C,o,a,b,c,d\}} \overrightarrow{\nabla_X(k)} . \overrightarrow{dX}$$

where $\overrightarrow{\nabla_X(k)}$ is the gradient of $k(\ldots, X, \ldots)$ with respect to point X. To give an expression of those gradients we need some preliminary results.

Let $\vec{v} = (v_x, v_y)^t$ be a vector in the reference plane in the scene, or in the image plane. Let us note $\vec{v}^\perp = (v_y, -v_x)^t$. Let $\vec{u}$ be another vector of the same plane. Then the gradient of the determinant $[\vec{u}, \vec{v}]$ is:

$$\overrightarrow{\nabla_u[\vec{u}, \vec{v}]} = \vec{v}^\perp$$

We get from this:

$$\overrightarrow{\nabla_A[\overrightarrow{OA}, \overrightarrow{OB}]} = \overrightarrow{OB}^\perp \quad ; \quad \overrightarrow{\nabla_B[\overrightarrow{OA}, \overrightarrow{OB}]} = \overrightarrow{AO}^\perp \quad ; \quad \overrightarrow{\nabla_O[\overrightarrow{OA}, \overrightarrow{OB}]} = \overrightarrow{BA}^\perp$$

See figure 2 for a geometrical interpretation of those vectors.

These expressions can then be used to compute the vectors $\overrightarrow{\nabla_X(k)}$. For example:

$$\overrightarrow{\nabla_O(k)} = k \left\{ \frac{\overrightarrow{CA}^\perp}{[\overrightarrow{OA}, \overrightarrow{OC}]} + \frac{\overrightarrow{DB}^\perp}{[\overrightarrow{OB}, \overrightarrow{OD}]} - \frac{\overrightarrow{CB}^\perp}{[\overrightarrow{OB}, \overrightarrow{OC}]} - \frac{\overrightarrow{DA}^\perp}{[\overrightarrow{OA}, \overrightarrow{OD}]} \right\}$$

$$\overrightarrow{\nabla_A(k)} = k \left\{ \frac{\overrightarrow{OC}^\perp}{[\overrightarrow{OA}, \overrightarrow{OC}]} - \frac{\overrightarrow{OD}^\perp}{[\overrightarrow{OA}, \overrightarrow{OD}]} \right\}$$

If we consider that uncertainty on X is the maximum error Δ_X that can occur, then the maximum error on k is:

$$\Delta_k = \sum_{X \in \{O,A,B,C,o,a,b,c,d\}} ||\overrightarrow{\nabla_X(k)}||.\Delta_X$$

where $||.||$ stands for the vector modulus. Let's note that Δ_k is the maximum error that may occur on k, and thus is a pessimistic estimation of the precision of the result.

As an alternative, we might consider that the uncertainty on X is the standard deviation σ_X of the distance between the true and estimated positions of X. As stated earlier, errors on

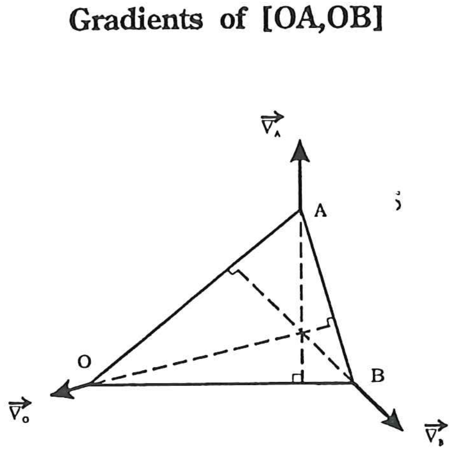

Figure 2: Gradients of $[\overrightarrow{OA}, \overrightarrow{OB}]$ with respect to the points O, A and B. Each gradient is orthogonal to the triangle side opposed to the vertex it is relative to.

the reference points are supposed to be small. If we can assume furthermore that these errors are independent, then the standard deviation of k is:

$$\sigma_k = \sqrt{\sum_{X \in \{O,A,B,C,o,a,b,c,d\}} \left(\sigma_X \, \overrightarrow{\nabla_X(k)}\right)^2}$$

Nervertheless, the hypothesis of independance of the reference points may be not valid if some of them result from a former localization process. For this reason, in the rest of this work, the uncertainty on k is taken as the maximum possible error Δ_k.

This value Δ_k is then classically used in conjunction with relation 3 to derive the maximum possible error on the direction (OD).

4 Minimizing the uncertainty

To identify a polygon in a perspective image, Sanfeliu *et al.* used their expression of the sensibility (5) to find out the combination of polygon vertices that produces the most precise invariant [5, 11, 12, 3].

Given four reference points O, A, B and C in a plane of the scene, and their respective image positions o, a, b and c, the above process allows to backproject on that plane any point d observed in the image. The result D is unique: it does not depend on the arbitrary order in which the reference points are considered.

Nevertheless if d happens to be aligned with two of the reference points, then one of the determinants that appears in expressions 1 and 4 may happens to be null; thus $\mathcal{I}$ and k may have a null numerator or denominator. If d is not aligned with two reference points but is close to one of the lines they form, then one or more determinant(s) may have a small absolute value; thus their relative uncertainty will be large, depending on the computer precision. It

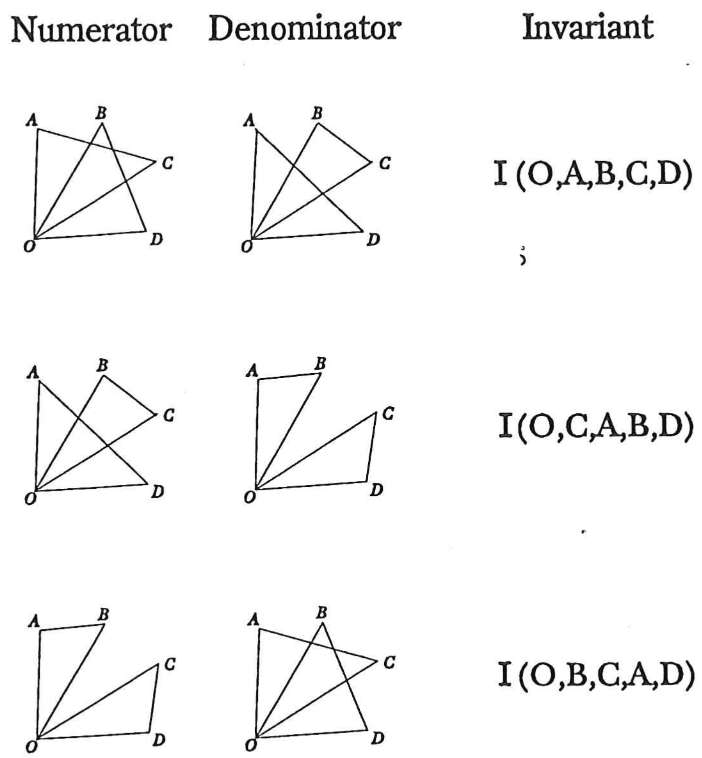

Figure 3: Three combinations of points A, B, C and D that produce expressions of the invariant $\mathcal{I}$ with the only three possible values of the relative uncertainties.

may be important to choose the role of the different reference points to avoid such situations, independently of any consideration on the precision of input data.

Furthermore, the above computation of the maximum possible error Δ_k on k, and of the uncertainty of the direction (OD) leads to a coarse majoration of the uncertainty that results from the unprecision of the input data. Computing with different orders of the reference points, thus, allows to keep only the smallest of those majorants.

Following this idea, the reference points A, B and C (and the image points a, b and c resp.) will be exchanged so that they give the best relative precision on the direction (OD). The projective invariant $\mathcal{I}$ is a combination of the signed areas of four triangles, all of them with the center point O as a vertice. The two triangles at numerator (resp. denominator) have no other vertice in common; the numerator and the denominator have no triangle in common; given the four points a, b, c and d, the choice of a triangle at numerator thus determines the other triangle at numerator. Considering that: 1) changing the order of triangles at numerator (resp. denominator) will not change the value of the invariant; 2) changing the order of vertices in a triangle may only change the sign of the invariant; 3) exchanging numerator and denominator would just give the inverse invariant; we finally obtain only three expressions of the invariant with different relative uncertainties (see figure 3). So it is necessary to compute the uncertainties only for those three configurations. Finally, the direction (OD) is computed using the most precise configuration.

This process is repeated with A, B, C considered as the center O' of a new pencil of lines defined by the other reference points. The direction of the line $(O'D)$ and its uncertainty are

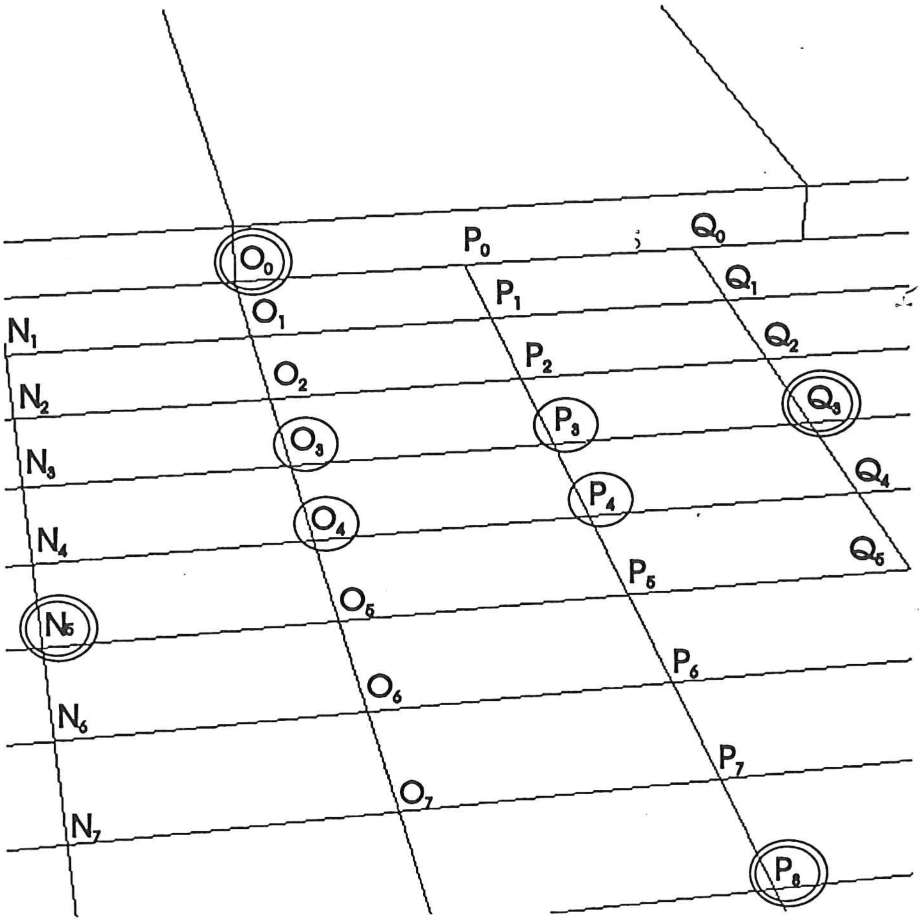

Figure 4: Synthesis perspective image of the test scene. The two sets of reference points that will be used are surrounded.

then computed.

Finally, all the pairs $\{(OD), (O'D)\}$ are considered; their intersections are computed, as well as their uncertainties. The intersection with the smallest uncertainty is retained as the resulting position of the scene point D.

5 Experimental results

This method has been implemented and used to compute the scene positions of 30 points of a synthetized image of a stone pavement (see figure 4). The scene is approximately 120 cm by 200 cm large, and the flagstones are 40 cm long and 25 cm wide, with an infinite precision. The image size is 512×512 pixels, and the image (reference and/or unknown) points are considered to be selected with a precision of 1 pixel.

		$N_\bullet$ $(-40.0, \bullet)$	$O_\bullet$ $(0.0, \bullet)$	$P_\bullet$ $(40.0, \bullet)$	$Q_\bullet$ $(80.0, \bullet)$
$\bullet_0$	$(\bullet, 800.0)$	–	$\boxed{(0.0, 800.0) \pm 1.9}$	$(40.0, 799.3) \pm 2.2$	$(79.9, 799.7) \pm 2.8$
$\bullet_1$	$(\bullet, 775.0)$	$(-40.2, 776.0) \pm 2.4$	$(-0.1, 775.5) \pm 1.7$	$(39.9, 775.9) \pm 2.0$	$(79.9, 775.0) \pm 2.3$
$\bullet_2$	$(\bullet, 750.0)$	$(-40.0, 750.4) \pm 2.2$	$(0.0, 750.8) \pm 1.8$	$(39.9, 750.0) \pm 1.7$	$(80.1, 749.9) \pm 2.0$
$\bullet_3$	$(\bullet, 725.0)$	$(-39.9, 725.8) \pm 1.7$	$(0.1, 725.6) \pm 1.5$	$(40.1, 724.9) \pm 1.5$	$\boxed{(80.0, 725.0) \pm 1.7}$
$\bullet_4$	$(\bullet, 700.0)$	$(-40.1, 700.0) \pm 1.5$	$(-0.1, 700.3) \pm 1.4$	$(40.0, 700.2) \pm 1.4$	$(80.0, 699.7) \pm 1.6$
$\bullet_5$	$(\bullet, 675.0)$	$\boxed{(-40.0, 675.0) \pm 1.3}$	$(-0.1, 675.2) \pm 1.3$	$(40.0, 675.1) \pm 1.3$	$(79.9, 675.6) \pm 1.6$
$\bullet_6$	$(\bullet, 650.0)$	$(-40.1, 649.7) \pm 1.5$	$(-0.1, 649.7) \pm 1.3$	$(40.2, 650.3) \pm 1.4$	–
$\bullet_7$	$(\bullet, 625.0)$	$(-40.2, 624.6) \pm 1.7$	$(-0.1, 625.1) \pm 1.4$	$(40.1, 625.0) \pm 1.2$	–
$\bullet_8$	$(\bullet, 600.0)$	–	–	$\boxed{(40.0, 600.0) \pm 1.1}$	–

Figure 5: Localization of the scene points shown in figure 4; the reference points that were used are shown in a box. The precision varies slowly between them.

Localizing the points in the scene was achieved with two sets of reference points. The points of the first set (namely O_0, N_5, Q_3 and P_8) are all around the scene, so that most of the unknown points are inside their quadrilateral. The other 26 points were localized with a computed uncertainty ranging from 1.1 to 2.8 cm. The maximum error observed is 1 cm (see figure 5).

	$N_\bullet$ $(-40.0, \bullet)$	$O_\bullet$ $(0.0, \bullet)$	$P_\bullet$ $(40.0, \bullet)$	$Q_\bullet$ $(80.0, \bullet)$
$\bullet_0$ $(\bullet, 800.0)$	–	$(-0.7, 798.9) \pm 13.6$	$(39.4, 800.9) \pm 14.1$	$(81.6, 804.0) \pm 26.9$
$\bullet_1$ $(\bullet, 775.0)$	$(-38.9, 773.1) \pm 16.1$	$(-0.7, 774.5) \pm 8.2$	$(39.5, 777.0) \pm 8.7$	$(81.6, 778.0) \pm 19.5$
$\bullet_2$ $(\bullet, 750.0)$	$(-38.4, 748.3) \pm 11.8$	$(-0.3, 750.0) \pm 4.1$	$(39.6, 750.5) \pm 4.2$	$(82.0, 751.9) \pm 14.3$
$\bullet_3$ $(\bullet, 725.0)$	$(-38.1, 724.5) \pm 9.7$	$\boxed{(0.0, 725.0) \pm 1.3}$	$\boxed{(40.0, 725.0) \pm 1.4}$	$(82.0, 725.8) \pm 11.3$
$\bullet_4$ $(\bullet, 700.0)$	$(-37.9, -40.1) \pm 9.6$	$\boxed{(0.0, 700.0) \pm 1.2}$	$\boxed{(40.0, 700.0) \pm 1.3}$	$(82.1, 699.5) \pm 11.3$
$\bullet_5$ $(\bullet, 675.0)$	$(-37.6, 675.6) \pm 11.8$	$(0.2, 675.3) \pm 3.7$	$(40.2, 674.5) \pm 3.9$	$(82.1, 674.5) \pm 13.2$
$\bullet_6$ $(\bullet, 650.0)$	$(-37.5, 651.3) \pm 16.1$	$(0.4, 650.2) \pm 7.7$	$(40.5, 649.5) \pm 8.0$	–
$\bullet_7$ $(\bullet, 625.0)$	$(-37.3, 627.2) \pm 22.2$	$(0.6, 626.0) \pm 12.8$	$(40.6, 624.1) \pm 13.5$	–
$\bullet_8$ $(\bullet, 600.0)$	–	–	$(40.7, 599.0) \pm 20.2$	–

Figure 6: Localization of the scene points shown in figure 4; the reference points that were used are shown in a box. Precision decreases rapidly as the distance to the reference points increases.

The points of the second set of reference points (namely O_3, O_4, P_3 and P_4) are all located near the center of the scene, so that all the unknown points are out of their rectangle. The other 26 points were localized with a computed uncertainty ranging from 1.2 to 26.9 cm. The maximum error observed is 4.3 cm (see figure 6).

6 Conclusions

A method was proposed to compute the precision of localization using projective invariants.

In recognition applications, this technique allows to choose rapidly the configuration of reference points that produces the smallest of the only three possible relative uncertainties.

Second, this method can be easily extended to compute the precision of volumetric invariants: triangles are replaced with tetrahedrons, and gradients of determinants of two vectors in the plans are replaced by gradients of determinants of three vectors in space, with an interpretation similar to that on figure 2.

At last, projective invariants appear to be more precise than stereovision [10]. Morin *et al.* point out that projective invariants (with infinitely precise reference points) provide the same precision as a perfectely calibrated camera.

This advantage on precision, plus the advantage of working with only one image, make projective invariants a valuable tool for urban cartography when the visible part of the scene contains enough reference points.

References

[1] Eamon B. Barrett, Paul M. Payton, Nils N. Haag, and Michael H. Brill. General methods for determining projective invariants in imagery. *CVGIP: Image Understanding*, 46–65, 1991.

[2] Hadj Batatia, Alain Ayache, and Charlie Krey. Reconnaissance d'objets polyédriques à partir d'une seule image, utilisant des invariants projectifs. In *Actes du* $8^{ème}$ *Congrès RFIA, Lyon*, pages 323–328, AFCET-INRIA, 1991.

[3] Walter Emde and Alberto Sanfeliu. Un método robusto para la extracción de invariantes proyectivos a través de los vértices de una proyección de un objeto 3D. In *Libro de comunicaciones, V Simposium National de Reconocimiento de Formas y Análisis de Imágenes, Valencia*, pages 39–46, 1992.

[4] Slimane Larabi. *Utilisation de la géométrie projective en vision stéréoscopique.* PhD thesis, INP Toulouse, 1991.

[5] Antoni Llorens Castelló. *Proyectividades e invariantes algebraicos en visión por computador.* PhD thesis, Universidad Politécnica de Cataluña, 1991.

[6] R. Mohr and L. Morin. Relative positioning from geometric invariants. In *Proc. CVPR'91, Hawai*, pages 139–144, IEEE, 1991.

[7] Luce Morin, Roger Mohr, and Enrico Grosso. Positionnement relatif à partir d'invariants projectifs. In *Actes du* $8^{ème}$ *Congrès RFIA, Lyon*, pages 311–321, AFCET-INRIA, 1991.

[8] Michael A. Penna and Richard R. Patterson. *Projective geometry and its applications to computer graphics.* Prentice Hall, 1986.

[9] J. A. Pérez, P. Horain, and J. J. Villanueva. CAVA: a system for Cartography Assisted by Vision Applications. 1992. Submitted to 8^{th} SCIA, Tromso, 1993.

[10] Josep Antoni Pérez López. *Automatización del proceso de adquisición de datos para la cartografía urbana usando técnicas de visión por computador.* Master's thesis, Univ. Autònoma Barcelona, 1992.

[11] A. Sanfeliu and A. Llorens. Projective invariants to identify polyhedric 3D objects. In N. Perez de la Blanca, A. Sanfeliu, and E. Vidal, editors, *Selected papers from the IVth Spanish Symposium (1990)*, pages 65–77, 1992.

[12] Alberto Sanfeliu, Antoni Llorens, and Walter Emde. Sensibility, relative error and error probability of projective invariants of plane surfaces of 3D objects. In *Proc. 11th IAPR, The Hagues*, pages 328–331, 1992.

[13] Jorge Stolfi. *Oriented projective geometry.* Academic Press, 1991.

[14] Roger Y. Tsai. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics and Automation*, 323–344, 1987.

Dépôt légal : 2ème trimestre 1993
Imprimé à l'ENST - PARIS
ISSN : 0751-1345

Ecole Nationale Supérieure des Télécommunications
46 rue Barrault 75634 Paris Cedex 13
Téléphone (1) 45 81 77 77 Télécopie (1) 45 89 79 06 Télex SUPTLCM 200160F

L'Ecole Nationale Supérieure des Télécommunications fait partie de l'enseignement supérieur de France Télécom