

A Localisation Method with an Omnidirectional Vision Sensor Using Projective Invariant

Bruno MARHIC, El Mustapha MOUADDIB⁺ IEEE member, and Claude PEGARD.

Université Picardie Jules Verne (UPJV)
Groupe de Recherche sur l'Analyse et la Commande des SYstèmes (GRACSY)
7 rue du MOULIN NEUF, 80 000 AMIENS, France
E-mail: Bruno.Marhic@sc.u-picardie.fr

Abstract

In this paper, we first present a short description of our conic sensor. Furthermore, we will pay special attention to the famous projective invariant, represented by the cross-ratio. We will end this paper by presenting a new 1D cross-ratio application that enables us to localise our mobile robot SARAH efficiently. With the method that we propose using, the camera does not need to be calibrated and this vision system, based on an omnidirectional sensor, is a major advantage for the navigation of an autonomous mobile robot.

1 Introduction

The principal task of computer vision is object or scene recognition. In most vision-based approaches for localisation, environment modelling is rendered complex as the geometric properties are not invariant to the projective transformation nor to the change of viewpoint [8]. Moreover, the matching phase is generally very complex and time consuming and most approaches require camera calibration.

Our previous dealings with an omnidirectional vision system are exposed in [13][14]. In [13], the approach is purely combinational and is not completely immune to the parasite straight lines of the image. We define parasite straight lines as all straight lines that do not correspond, *a priori*, to a recorded reference point on the map. In [14], a solution to the problem of parasite straight lines is suggested by fusing the odometric sensors with the omnidirectional vision. The method suggested in [14] renders good results, but does necessitate a fusion of the sensors. In a recent work [15], K.S. Roh *et al* use projective invariants to localise a mobile robot. However, their method requires a vanishing point formed by the image of a corridor in order to be able to calculate an invariant in the scene.

In this paper, we present a new application of 1D cross-ratio (and its dual for four lines), in order to resolve the matching between the model (in any type of environment) and an omnidirectional image of the scene. Our matching method is trivial and does not require

camera calibration as we also use geometric projective invariants [11].

What are invariants? The idea of invariance in computer vision appears when one has analysed the human ability to recognise objects. Invariants are quantities that allow for matching an object (or a scene) in an image with a model, no matter the transformation applied to the model and no matter the viewpoint. We use the term "transformation" in its larger sense, meaning that it can contain the composition of several elementary linear transformations (rotation, translation, scaling, shearing...). More the transformation is general, complex, more it is difficult to find invariants. This leads us to the unavoidable question: for which type of geometric transformation do we want to calculate an invariant quantity?

The major part of invariants listed in the computer vision literature are *intrinsic invariants*, more often called *geometric descriptors* or *geometric invariants*; however there are also non-geometric invariants called *extrinsic invariants* [2].

We can distinguish two significant classes of invariants: *global invariants* (or *integrals*) [3] and *local invariants* (or *differentials*) [4]. A global invariant defines information on the object in its entirety, which renders it quite insensitive to noise. Such invariants necessitate knowledge of the entire object or scene, but are rather sensitive to occlusions. A local invariant defines information on a point or a small unity of points (*semi-local*) independently of the other points that constitute the contour of the object. These invariants (local and semi-local) are much more immunised to occlusions as the contour is not used in its entirety, but are rather sensitive to noises, which implies that in most cases a preliminary image filtering is needed. Generally, differential invariants methods require high orders of derivatives for calculating invariants, which is not without consequences for the reliability of these methods.

Eventhough all invariant quantities can normally be classed in either the global/integral invariant family, or in the local/differential invariant family, the *algebraic invariants* merit a new class. Algebraic methods imply the calculation of *implicit polynomes* of the contour [5], such

⁺ PhD Tutor of B. Marhic.

as conics, quadrics, etc. Therefore, contrary to differential invariants that are defined locally (a point), algebraic invariants remain constant for the entire implicit polynome considered; algebraic invariants are *regional invariants* (between local and global) and are much less sensitive to image noises compared to differential invariants. Examples of algebraic invariants are given in [6].

Using *quasi-invariants*, that are easier to obtain than invariants, might prove helpful. Unfortunately, these quantities (quasi-invariant) are only "invariant" on the limited transformation range for which they were calculated.

2 Intrinsic properties of our conic sensor: SYCLOP¹

Our conic sensor is made up of two essential parts: a CCD camera and a conic reflector (Figure 1-a). These two parts are separated by a glass support. This type of perception system is also used by YAGI [1] with the COPIS system. An example of an image generated by the omnidirectional sensor is given in Figure 1-b.

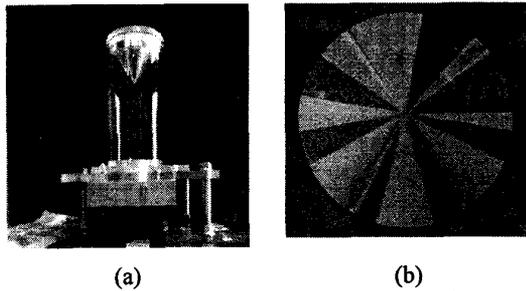


Figure 1: The omnidirectional sensor and a conic image.

Advantage for navigation: Associating a conic mirror with a camera enables us to obtain a panoramic image (360°) of the robot's environment. This characteristic is a major asset to the navigation of an autonomous mobile robot, as it takes away the robot's "direction of the vision field" constraint for a given robot's configuration $X_r^T = [x_r, y_r, \theta_r]_{Re}$, where Re is the absolute reference of the environment. This property of omnidirectionality offers a better chance for the matching phase between models and the scene and, therefore, a higher chance of localising itself according to the landmarks of the environment.

Geometrical advantage: The use of this particular sensor enables us to modelise the 3D environment of the robot into a simple 2D map, containing reference points expressed in the frame Re . These points represent the vertical landmarks of the environment. The orthographic projection of the vertex S of the cone (which lies on the revolution axis Δ), gives the point O in the image plane (π). This point O is materialised through the intersection of the virtual view lines in the image. Therefore, this point O is a virtual projection centre. As we can see from Figure

2, the line L' in the image plane (π) is the projective line of L with point O as the centre of projection; this geometric transformation is the fundamental 1D projective transformation. The 1D projective mapping between lines is given by the 2x2 homogeneous transformation matrix T , so that $x = T.X$ where $[x_1, x_2]^T$ are the homogeneous coordinates on the transformed line. There are three essential parameters to define the matrix T (over scale of T is not important).

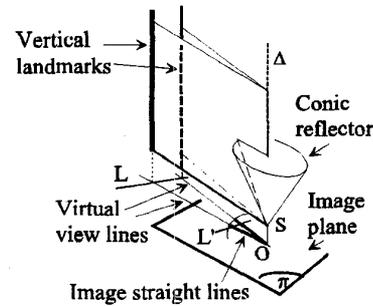


Figure 2: Radial straight lines formation and the 1D projective transformation.

3 Theoretical Preliminaries

3.1 Basic theory of cross-ratio

Cross-ratio (also known as anharmonic-ratio) has been recognised for a long time as being the most fundamental projective invariant. All other projective invariants can be derived from the cross-ratio. We would like to give one historical example. In 1827, the mathematician AF Möbius [7] proved the invariance of 1D cross-ratio and these generalisations to triangle areas and tetrahedric volumes. Chapter 11 of [8] presents, in a clear and relatively concise fashion, the cross-ratio for its use in computer vision.

Let $P_{1, \dots, 4}$ be any four collinear points in the plane. The following theorem [9], defines the cross-ratio:

Theorem: The cross-ratio of distances between any four points in the object line is the same as the cross-ratio of distances between their images in any image line (Figure 3).

Let ρ be the cross-ratio of the 4 collinear points:

$$CR(P_1, P_2, P_3, P_4) = \rho = \frac{D'_{13}D'_{24}}{D'_{14}D'_{23}} = \frac{D_{13}D_{24}}{D_{14}D_{23}}, \quad (1)$$

where D_{ij} is the distance from $P_i = [x_i, y_i]^T$ to $P_j = [x_j, y_j]^T$ along the line L (*idem* for D'_{ij}). The cross-ratio is an absolute invariant. Proof of the preceding theorem is given in [9].

¹ SYstème Conique pour la LOcalisation et la Perception.

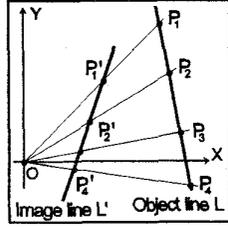


Figure 3: Theorem for point projections and cross-ratios.

Fundamental theorem [11]: Any homography preserves the cross-ratio.

A homographic transformation is any linear transformation in homogeneous co-ordinates, including central projection (perspective projection), linear scalings, skewings, rotations, translations ... Thus, our conic sensor preserves the cross-ratio.

Since points and lines are dual, an equivalent cross-ratio for lines exists. The dual relation to collinearity is incidence at a point. A cross-ratio is defined on four lines which are incident at a single point. Any set of lines incident at a common point is called a pencil. The cross-ratio of a pencil can be defined in terms of the angles between the lines as shown in Figure 4, and is given by:

$$CR(L_1, L_2, L_3, L_4) = \frac{\sin \alpha_{13} \sin \alpha_{24}}{\sin \alpha_{23} \sin \alpha_{14}} = \rho \quad (1bis)$$

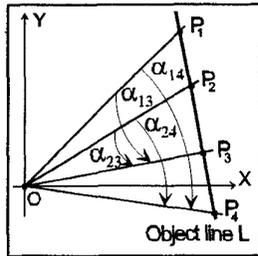


Figure 4: The dual configuration for the cross-ratio.

Unfortunately, the cross-ratio depends on the order in which the points are marked. If the labels of the four points are permuted, the cross-ratio of four points renders only six different values (of which three are the inverse of the other three) for twenty-four (4!) possible labellisations. The six different values in relation to ρ are [10]:

$$\begin{aligned} \rho_1 &= \rho & \rho_3 &= 1 - \rho & \rho_5 &= \rho_1(-\rho_4) \\ \rho_2 &= \rho_1^{-1} & \rho_4 &= \rho_3^{-1} & \rho_6 &= \rho_2(-\rho_3) \end{aligned}$$

In order to remedy this numeration problem (that can hinder the recognition of objects), we can use symmetric functions of ρ , that are invariant to the permutation of labels [11][12]. The functions $I_1(\rho)$, $I_2(\rho)$, and $I_3(\rho)$ are examples of permutation invariant functions:

$$I_1(\rho) = \sum_{i=1, \dots, 6} \rho_i, \quad (2)$$

$$I_2(\rho) = \prod_{i=0, \dots, 2} (\rho_j + \rho_{j+1}) \quad \text{with } j = 2i + 1, \quad (3)$$

$$I_3(\rho) = \frac{(\rho^2 - \rho + 1)^3}{\rho^2(\rho - 1)^2}, \quad (4)$$

where $I_3(\rho)$ is called the j-invariant. Unfortunately, the use of such functions (I_1 and I_2) of ρ is very dainty as these functions are unstable. As a matter of fact, only one cross-ratio $\rho_{i=1, \dots, 6}$ needs to be unstable to render the functions $I_{k=1,2}(\rho)$ unstable as well. The use of the j-invariant (eq. (4)) is preferable.

3.2 Calculation of the Position

In the static case, the robot's configuration is given by the vector X:

$$X = [x_c \quad y_c \quad \theta]_{Re},$$

where (x_c, y_c) are the robot's co-ordinates in the world frame Re and θ is the robot's orientation in the same frame. The robot's configuration is determined by the equations linking the sensorial radial straight lines (in the image) with the co-ordinates of the landmarks (eq. (5)).

$$\tan(\theta + \phi_i) = \frac{y_c - y_i}{x_c - x_i}, \quad (5)$$

where ϕ_i is the relative angle of the landmark B_i in the robot's frame R_r and (x_i, y_i) are the co-ordinates of the landmark B_i . We obtain one equation for each matched radial straight line. The system of non-linear equation obtained is identical to the one managed by the goniometrical localisation methods [16]. We refer the interested readers to [13][14], where this localisation method is treated.

4 Experimental Results

On the basis of the discussion above, our experimental results with real images are shown. We have implemented our localisation algorithm on the mobile robot SARAH.

The map of the environment is known *a priori*. The environment's map creation is an off-line process. This map is simply a set of reference points in the world's reference frame Re. These points represent the absolute location of the vertical landmarks (doors, wall's junctions...).

4.1 Candidates' formation for matching

In order to simplify the following experimental results, we voluntarily limited ourselves to determining just one invariant-model with two aligned and contiguous doors of our laboratory (Figure 5). It is easy to generalise the following principle for each quadruplet of collinear points of the environment's map and design a model table.

Invariant-model design: As we mentioned above, we calculate an invariant-model pair with two doors. We simply apply the equations (1) and (4), where $D_{13} = \text{dist}(P_1, P_3) = 117.7 \text{ Cm}$, $D_{14} = 198 \text{ Cm}$, $D_{24} = 107.5 \text{ Cm}$ et $D_{23} = 27.2 \text{ Cm}$. The value of the model-invariant is given by the $(\rho_{\text{mod}}, j_{\text{mod}})$ pair (eq. (6)). The evaluation of the model-invariant is an off-line calculation.

$$\begin{cases} \rho_{\text{mod}} = (117.7/198)/(27.2/107.5) = 2.3494 \\ j_{\text{mod}} = 7.2161 \end{cases} \quad (6)$$

Primitive extraction: For each image, we first apply a low pass filter of a 3x3 weighted kernel window in order to reduce the noise. We then calculate a 3x3 Sobel filter in order to extract the boundaries of the conic image (Figure 6). Finally, we apply a modified Hough transformation in order to design the conic image's radial straight lines (Figure 7).

Candidates formation: Let n be the number of primitives extracted from the conic image. Then, there are $C(n,4)$ combinations of quadruplets of straight lines. We calculate one cross-ratio for each quadruplet and compare it with the model. As the cross-ratio is not a metric dimension, a comparison by difference makes no sense. We have to implement the distribution function $F(\rho)$ of the cross-ratio to be able to compare the different cross-ratios. Thus, we take the following equation [17] as the distance between two cross-ratios:

$$d(\rho_1, \rho_2) = \min(|F(\rho_1) - F(\rho_2)|, 1 - |F(\rho_1) - F(\rho_2)|),$$

where F is the distribution function. The projective distance d will be between 0 and $\frac{1}{2}$. Then, we reject all cross-ratios too far from the model. We also apply two trivial criteria to decrease the number of candidates. First, as the model is calculated from four aligned points on the same wall, we can reject, without any loss of information, the angular sectors (formed by the lines of the quadruplet) bigger than 180 degrees. Secondly, we can also reject all the angular sectors smaller than 10 degrees, this means that the robot is too far from the four landmarks. All the non-rejected quadruplets are candidates for the matching/localisation step.

The following table, the Table 1, shows the effects of the different criteria on the numbers of candidates for the images (a)...(d) of the Figure 7.

Table 2 shows the values of invariants for the quadruplet of straight lines in the image that correspond with the vertical landmarks that have served to determine the model-invariant. There is an invariant pair for each image (a), ...(d) of Figure 6. In Table 2, we have also shown the mean m over the four images, the standard deviation σ and the percentage of the standard deviation to the mean. Table 2 shows that, essentially, all of the values ρ_i and j_i are constants; they remain stable in spite of the change in viewpoint. This shows that the values of the invariant are reliable, even for noisy images. One can also

notice that invariants $(\rho_{\text{mod}}, j_{\text{mod}})$ are between $m_{\rho, j} \pm \sigma_{\rho, j}$; this shows that 1D cross-ratio is reliable, when one considers projective transformation.

Rejection criteria	Nb of Candidates			
	(a)	(b)	(c)	(d)
$C(n,4)$ quadruplets	7315	3060	7315	7315
Angular sectors $>180^\circ$	2114	1115	1381	2225
Angular sectors $<10^\circ$	2108	1115	1380	2219
$d(\rho_{\text{mod}}, \rho_i) > 0.006$	175	63	58	105

Table 1: Number of candidates after the different criteria.

	Cross Ratio	J-invariant
(a)	2.4285	7.4170
(b)	2.2161	6.9459
(c)	2.2620	7.0283
(d)	2.3650	7.2535
Model	2.3494	7.2161
m (Mean)	2.3179	7.1612
σ (Std. Dev.)	0.0965	0.2145
σ/m (%)	4.16	3.00

Table 2: Values of invariants.

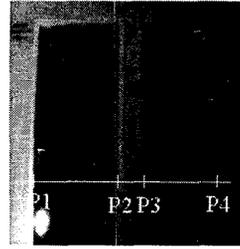


Figure 5: Geometric configuration for the model.

4.2 Matching and localisation

The use of invariants makes the matching phase trivial and efficient. We calculate a hypothetic configuration for each selected quadruplet in the image. In our case, the system of equations (5) to be resolved is overdetermined (four lines). An iterative method, such as the Newton-Rapson's, could be used here. Then, to validate the hypothetic configuration, we try to match all other landmarks of the map with the lines in the image. Just one quadruplet of lines in the image has to verify the previous condition.

Table 3 shows the localisation results from the images of the Figure 7. The configuration $X = [x_c \ y_c \ \theta]_{Re}$ is expressed in millimetres for $(x_c; y_c)$ and in degrees for θ .

Images	Measured location	Calculated location
(a)	[1680; 1472; 340]	[1641; 1532; 345]
(b)	[1470; 2020; 0]	[1493; 2045; 354]
(c)	[2500; 1020; 300]	[2547; 994; 306]
(d)	[3775; 2775; 310]	[3751; 2808; 313]

Table 3: Localisation results.

The experimental results shown in Table 3 are relatively satisfying but, unfortunately, they are not accurate enough. However, we can easily increase the precision, if we pay more attention to the different steps of this localisation treatment. The errors come from: model-map (environment) design, angles estimation (round to degree),...

5 Conclusions

In this paper, we have first presented a brief overview of the different techniques of invariants used in computer vision. We then concentrated on the famous projective invariant, which is the cross-ratio. Next, with the help of an application using a conic sensor, we demonstrated that the cross-ratio is reliable and suitable when the projective transformations and real noisy images are taken into account. We have also shown that geometric invariants are useful tools for the localisation of a mobile robot, because they are viewpoint invariant.

In the case where one considers more than one invariant pair as model to define the environment, occlusion of any vertical landmark does not pose a problem because our matching method is semi-local.

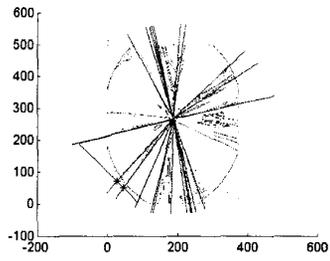
Man-made map design is a fastidious and cumbersome task, thus we now work to substitute this map by real conic images as model base.

Acknowledgement

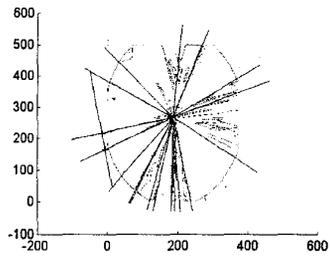
This work was supported in part by the "Conseil Régional de Picardie" under the project "Pôle DIVA".

References

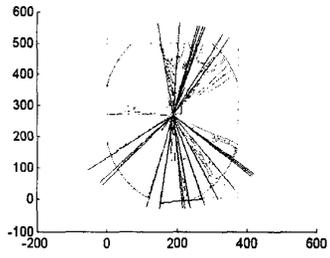
- [1] Y. Yagi, Y. Nishizawa and M. Yachida / Map-based Navigation for a Mobile Robot with Omnidirectional Image Sensor COPIS / IEEE Trans. on Robotics and Automation, Vol RA-11, pp 634-648, October 1995.
- [2] C. Schmid and R. Mohr / Combining Grey-value Invariants with Local Constraints for Object Recognition / In Proc. Computer Vision and Pattern Recognition, CVPR, pp872-877, 1996.
- [3] M. K. Hu / Visual Pattern Recognition by Moment Invariants / IRE Trans. Inform Theory, Vol IT-8, pp 179-187, February 1962.
- [4] I. Weiss / Projective Invariants of Shapes / In Proc. Computer Vision and Pattern Recognition, CVPR, pp 291-297, 1988.
- [5] A. Shashua / Algebraic Functions for Recognition / IEEE Trans. Pattern Analysis and Machine Intelligence, Vol PAMI-17, No 8, pp 779-789, August 1995.
- [6] D. Keren / Using Symbolic Computation to Find Algebraic Invariants / IEEE Trans. Pattern Analysis and Machine Intelligence, Vol PAMI-16, No 11, pp 1143-1149, November 1994.
- [7] A. F. Möbius / Der Barycentrische Calcul / Verlag von Johann Ambrosius Barth, Leipzig, 1827; in A. F. Möbius, "Gesammelte-werke, Vol 1", Dr Martin Sandig, OHG, Wiebaden, 1967.
- [8] R. O. Duda and P. E. Hart / Pattern Classification and Scene Analysis / New York: Wiley 1973.
- [9] E. B. Barrett, P. M. Payton, N. N. Haag and M. H. Brill / General Methods for Determining Projective Invariants in Imagery / CVGIP: Image Understanding, Vol CVGIP:IU-53, No 1, pp 46-65, January 1991.
- [10] S. J. Maybank / Classification on the Cross Ratio / Applications of Invariance in Computer Vision, Lecture Notes in Computer Science, LNCS-825, pp 453-472, 1993.
- [11] J.L. Mundy, A. Zisserman / Geometric Invariance in Computer Vision / Cambridge, Mass.:MIT Press, 1992
- [12] T. H. Reiss / Recognizing Planar Objects Using Invariant Image Features / Lectures Notes in Computer Vision, LNCS-676, Springer-Verlag, 1992.
- [13] C. Pégard and M. Mouaddib / A Mobile Robot Using a Panoramic View / In Proc. IEEE Int. Conf. on Robotics and Automation, ICRA, pp 89-94, 1996.
- [14] L. Delahoche, C. Pégard, B. Marhic and P. Vasseur / A Navigation System on an Omnidirectional Vision Sensor / In Proc. IEEE/RSJ Int. Conf. on Intelligent RObots and Systems, Vol IROS-2, pp718-724, 1997.
- [15] K. S. Roh, W. H. Lee, and I. S. Kweon / Obstacle Detection and Self-localisation without Camera Calibration Using Projective Invariants / In Proc. IEEE/RSJ Int. Conf. on Intelligent RObots and Systems, Vol IROS-2, pp1030-1035, 1997.
- [16] Durant-Whyte / The Design of a Radar-based Navigation System for Large Outdoor Vehicles / In Proc. IEEE Int. Conf. on Robotics and Automation, ICRA, pp 764-769, May 1995.
- [17] P. Gros / Outils Géométriques pour la Modélisation et la Reconnaissance d'Objets Polyédriques / Thèse INRIA Rhône-Alpes, spécialité Informatique, Juillet 1993.



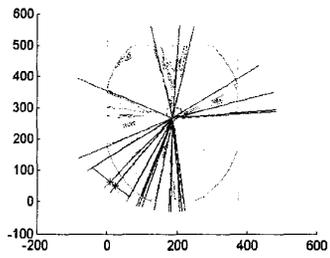
(a)



(b)

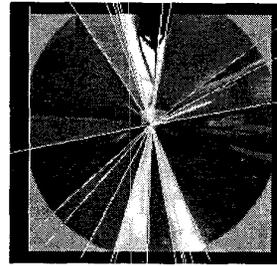


(c)

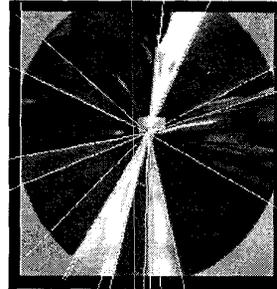


(d)

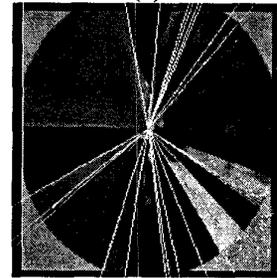
Figure 6: Extracted straight lines from conic images and cross-ratio determination.



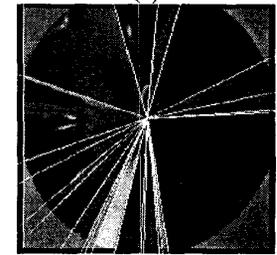
(a)



(b)



(c)



(d)

Figure 7. Straight lines and conic images.