



Homography-based Tracking for Central Catadioptric Cameras

Christopher Mei, Selim Benhimane, Ezio Malis, Patrick Rives

► To cite this version:

Christopher Mei, Selim Benhimane, Ezio Malis, Patrick Rives. Homography-based Tracking for Central Catadioptric Cameras. 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, Oct 2006, Beijing, China. pp.669-674, 10.1109/IROS.2006.282553 . hal-04654776

HAL Id: hal-04654776

<https://hal.science/hal-04654776v1>

Submitted on 20 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Homography-based Tracking for Central Catadioptric Cameras

Christopher MEI, Selim BENHIMANE, Ezio MALIS and Patrick RIVES

INRIA Icare Project-team

Sophia-Antipolis, France

Email : {firstname.surname@sophia.inria.fr}

Abstract—This paper presents a parametric approach for tracking piecewise planar scenes with central catadioptric cameras (including perspective cameras). We extend the standard notion of homography to this wider range of devices through the unified projection model on the sphere. We avoid unwarping the image to a perspective view and take into account the non-uniform pixel resolution specific to non-perspective central catadioptric sensors. The homography is parametrised by the Lie algebra of the special linear group $\mathbb{SL}(3)$ to ensure that only eight free parameters are estimated. With this model, we use an efficient second-order minimisation technique leading to a fast tracking algorithm with a complexity similar to a first-order approach. The developed algorithm was tested on the estimation of the displacement of a mobile robot in a real application and proved to be very precise.

I. INTRODUCTION

The wide field of view offered by panoramic cameras make them ideal candidates in tasks such as motion estimation, autonomous navigation and localization [1]. Catadioptric cameras are the omnidirectional vision systems that combine a single mirror and a camera. Baker and Nayar [2] derive the complete class of catadioptric sensors comprised of a unique mirror and with a single effective viewpoint, these are referred to as *central* catadioptric sensors. Geyer [3] and Barreto [4] developed a unified projection model for these sensors using properties of the projection of points on the sphere. It has since been used extensively to develop generic algorithms. Recently a lot of research has been done in ego-motion estimation [5], [6] and visual servoing [7], [8]. Tracking, which is a fundamental step for various computer vision applications, has seen very few articles for catadioptric systems, [7] being one of the rare examples. This approach however was only applied to track straight lines using a contour-to-point tracker to avoid the problem of quadric-based catadioptric line fitting. The non-linear projection model of the sensor means that even simple motions such as translations modify the shape to track significantly. Unwarping the image to a perspective view introduces discontinuities in the Jacobian (at least two planes are needed to represent the 360° field of view), is inefficient and means the non-uniform pixel resolution of the sensor is not taken into account. This leads naturally to the use of parametric models [9], [10], [11], [12] such as the homography-based approach presented in this article. Previous related work for perspective cameras were done in [13] which extends the work proposed by Hager [10] using

homographies and in [14]. Homographies have also been used recently for visual servoing with central catadioptric cameras [8] and share with this article the notion of homographies for points belonging to the sphere of the unified projection model.

The tracking approach presented here minimizes a dissimilarity measure and more specifically a sum-of-squared-differences (SSD) between a reference template and the current image warped through a homography. This leads to a non-linear optimization problem that can be solved for small displacements (the type of movement that would be expected in a scene at video rate). Compared to methods such as [15] we do not linearize the state estimate but obtain the optimal transformation at each new incoming image. This makes the approach less sensitive to strong camera motion. The rotation and the translation of the camera can then be extracted from the obtained optimal homography [16]. The main contribution of this article concerns the extension of homography-based tracking to central catadioptric devices. We will see that through the projection model we take into account the nature of the sensor (non-uniform resolution and distortion). We apply the efficient second-order minimization algorithm proposed by Malis [14], [17] in order to improve the convergence domain and speed of standard first-order minimization algorithms [9], [10], [11], [12]. The algorithm is evaluated on the motion estimation of a mobile robot and the results are compared to the precise odometry considered as ground truth.

II. LIE-GROUP HOMOGRAPHY PARAMETERIZATION

A. The $\mathbb{SL}(3)$ Lie Group

Let $\mathbf{R} \in \mathbb{SO}(3)$ be the rotation of the camera and $\mathbf{t} \in \mathbb{R}^3$ its translation. \mathbf{R} can be written as $\mathbf{R} = \exp([\mathbf{r}]_\times)$ where $\mathbf{r} = \mathbf{u}\theta$ (\mathbf{u} and θ being the axis and the angle of rotation respectively). The standard planar homography matrix \mathbf{H} is defined up to a scale factor:

$$\mathbf{H} \sim \mathbf{R} + \mathbf{t}\mathbf{n}_d^{*\top}$$

where $\mathbf{n}_d^* = \mathbf{n}^*/d^*$ is the ratio between the normal vector to the plane \mathbf{n}^* and the distance d^* of the plane to the origin of the reference frame. In order to fix the scale, we choose $\det(\mathbf{H}) = 1$, i.e. $\mathbf{H} \in \mathbb{SL}(3)$ (the Special Linear group of dimension 3). This choice is well justified since $\det(\mathbf{H}) = 0$ happens only if the observed plane passes through the optical center of the camera (in this singular case the plane is not

visible any more). Let $\mathbf{w} : \text{SL}(3) \times \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be an action of $\text{SL}(3)$ on \mathbb{R}^2 on the left (i.e. $\mathbf{w}\langle\mathbf{H}\rangle\langle\mathbf{p}\rangle \in \mathbb{R}^2$). The map $\mathbf{w}\langle\mathbf{H}\rangle : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ defines a coordinate transformation (a warping) such that $\mathbf{q} = \mathbf{w}\langle\mathbf{H}\rangle\langle\mathbf{p}\rangle$. Let \mathbf{I} be the identity element of the transformation group :

- $\mathbf{w}\langle\mathbf{I}\rangle\langle\mathbf{p}\rangle$ is the identity map. i.e. $\forall \mathbf{p} \in \mathbb{R}^2$:

$$\mathbf{w}\langle\mathbf{I}\rangle\langle\mathbf{p}\rangle = \mathbf{p} \quad (1)$$

- the composition of two actions corresponds to the action of the composition, i.e. $\forall \mathbf{p} \in \mathbb{R}^2, \forall \mathbf{H}_1, \mathbf{H}_2 \in \text{SL}(3)$:

$$\mathbf{w}\langle\mathbf{H}_1\rangle\langle\mathbf{w}\langle\mathbf{H}_2\rangle\langle\mathbf{p}\rangle\rangle = \mathbf{w}\langle\mathbf{H}_1\mathbf{H}_2\rangle\langle\mathbf{p}\rangle \quad (2)$$

B. The Lie algebra of $\text{SL}(3)$ and the exponential map

Let \mathbf{A}_i , with $i \in \{1, 2, \dots, 8\}$, be a basis of the Lie algebra $\mathfrak{sl}(3)$ (i.e. the dimension of $\mathfrak{sl}(3)$ is 8). Any matrix $\mathbf{A} \in \mathfrak{sl}(3)$ can be written as a linear combination of the matrices \mathbf{A}_i :

$$\mathbf{A}(\mathbf{x}) = \sum_{i=1}^8 x_i \mathbf{A}_i \quad (3)$$

where $\mathbf{x} = (x_1, x_2, \dots, x_8)$ is a (8×1) vector and x_i is the i -th element of the base field. The \mathbf{A}_i matrices are defined as:

$$\begin{aligned} \mathbf{A}_1 &= \begin{bmatrix} 001 \\ 000 \\ 000 \end{bmatrix}, \mathbf{A}_3 = \begin{bmatrix} 010 \\ 000 \\ 000 \end{bmatrix}, \mathbf{A}_5 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \mathbf{A}_7 = \begin{bmatrix} 000 \\ 000 \\ 100 \end{bmatrix} \\ \mathbf{A}_2 &= \begin{bmatrix} 000 \\ 001 \\ 000 \end{bmatrix}, \mathbf{A}_4 = \begin{bmatrix} 000 \\ 100 \\ 000 \end{bmatrix}, \mathbf{A}_6 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{A}_8 = \begin{bmatrix} 000 \\ 000 \\ 010 \end{bmatrix} \end{aligned} \quad (4)$$

The exponential map links the Lie algebra to the Lie Group: $\exp : \mathfrak{sl}(3) \rightarrow \text{SL}(3)$. There exist an open cube v about $\mathbf{0}$ in $\mathfrak{sl}(3)$ and an open neighborhood U of the identity matrix \mathbf{I} in $\text{SL}(3)$ such that $\exp : v \rightarrow U$ is smooth and one-to-one onto, with a smooth inverse. The neighborhood U of \mathbf{I} is very large. Consequently, the homography matrix \mathbf{H} can be locally parameterized as:

$$\mathbf{H}(\mathbf{x}) = \exp \left(\sum_{i=1}^8 x_i \mathbf{A}_i \right)$$

III. GENERALIZED HOMOGRAPHY-BASED TRACKING

A. Unified projection model

For sake of completeness, we present here a slightly modified version of the projection model of Geyer [3] and Barreto [4] (Fig. 1). The projection of 3D points can be done in the following steps (the values for ξ and γ are detailed Table I):

- 1) world points in the mirror frame are projected onto the unit sphere, $(\mathcal{X})_{\mathcal{F}_m} \rightarrow (\mathcal{X}_s)_{\mathcal{F}_m} = \frac{\mathcal{X}}{\|\mathcal{X}\|} = (X_s, Y_s, Z_s)$
- 2) the points are then changed to a new reference frame centered in $\mathbf{C}_p = (0, 0, \xi)$, $(\mathcal{X}_s)_{\mathcal{F}_m} \rightarrow (\mathcal{X}_s)_{\mathcal{F}_p} = (X_s, Y_s, Z_s - \xi)$
- 3) we then project the point onto the normalized plane, $\mathbf{m} = (\frac{X_s}{Z_s - \xi}, \frac{Y_s}{Z_s - \xi}, 1) = \mathbf{h}(\mathcal{X}_s)$

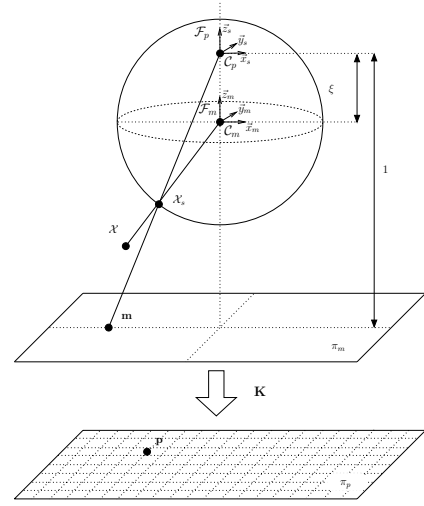


Fig. 1. Unified image formation

- 4) the final projection involves a generalized camera projection matrix \mathbf{K} (with γ the generalized focal length, (u_0, v_0) the principal point, s the skew and r the aspect ratio)

$$\mathbf{p} = \mathbf{K}\mathbf{m} = \begin{bmatrix} \gamma & \gamma s & u_0 \\ 0 & \gamma r & v_0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{m} = k(\mathbf{m})$$

The function \mathbf{h} is bijective and

$$\mathbf{h}^{-1}(\mathbf{m}) = \begin{bmatrix} \frac{-\xi - \sqrt{1 + (1 - \xi^2)(x^2 + y^2)}}{x^2 + y^2 + 1} x \\ \frac{-\xi - \sqrt{1 + (1 - \xi^2)(x^2 + y^2)}}{x^2 + y^2 + 1} y \\ \frac{-\xi - \sqrt{1 + (1 - \xi^2)(x^2 + y^2)}}{x^2 + y^2 + 1} + \xi \end{bmatrix} \quad (5)$$

We will call lifting the calculation of the point \mathcal{X}_s corresponding to a given point \mathbf{m} (or \mathbf{p} according to the context). We may note that in the perspective case, there is no mirror and only points with $Z > 0$ are accepted (we thus fall back to the standard projection model with an extra normalization to the sphere).

TABLE I
UNIFIED MODEL PARAMETERS

	ξ	γ
Parabola	1	$-2pf$
Hyperbola	$\frac{df}{\sqrt{d^2 + 4p^2}}$	$\frac{-2pf}{\sqrt{d^2 + 4p^2}}$
Ellipse	$\frac{df}{\sqrt{d^2 + 4p^2}}$	$\frac{2pf}{\sqrt{d^2 + 4p^2}}$
Planar	0	$-f$
Perspective	0	f
d : distance between focal points $4p$: latus rectum		

B. Minimization problem

Let I^* be the reference image. We will call reference template, a region of size q (rows \times columns) of I^* corresponding to the projection of a planar region of the scene.

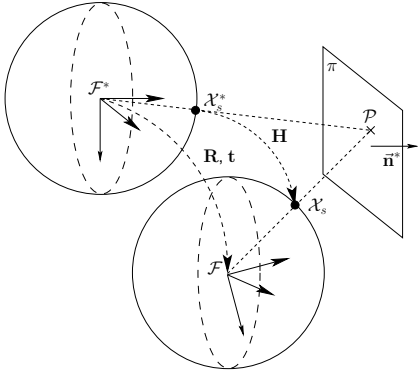


Fig. 2. Homography between points on the sphere

Two planar points are related by a homography \mathbf{H} by $\mathcal{X} = \mathbf{H}\mathcal{X}^*$ so the projection of points, belonging to a planar region of the scene, on the sphere are related by the planar homography \mathbf{H} : $\rho\mathcal{X}_s = \rho^*\mathbf{H}\mathcal{X}_s^*$.

Consider the following scheme, illustrated by figure 2 (n corresponds to the point normalization):

$$\begin{array}{ccccccc} \mathbf{p}^* & \xrightarrow{k^{-1}} & \mathbf{m}^* & \xrightarrow{h^{-1}} & \mathcal{X}_s^* & & \\ & & & & \downarrow \mathbf{H} & & \\ \mathbf{p} & \xleftarrow{k} & \mathbf{m} & \xleftarrow{h} & \mathcal{X}_s & \xleftarrow{n} & \mathcal{X} \end{array} \quad (6)$$

Let $\Pi = k \circ h$ be the transformation between the sphere and the image plane: $\mathbf{p} = \Pi(\mathcal{X}_s)$. To track the template in the current image I is to find the transformation $\bar{\mathbf{H}} \in \mathbb{SL}(3)$ that warps the lifting of that region to the lifting of the reference template of I^* , i.e. find $\bar{\mathbf{H}}$ such that:

$$\forall i : I(\Pi(n(\mathbf{w} \langle \bar{\mathbf{H}} \rangle \langle \mathcal{X}_s^{i*} \rangle))) = I^*(\mathbf{p}_i)$$

In other words, knowing an approximation $\hat{\mathbf{H}}$ of the transformation $\bar{\mathbf{H}}$, the problem is to find the incremental transformation $\mathbf{H}(\mathbf{x})$ that minimizes the sum of squared differences (SSD) over all the pixels:

$$\begin{cases} F(\mathbf{x}) = \frac{1}{2} \sum_{i=1}^q \|\mathbf{f}_i\|^2 \\ \mathbf{f}_i = I\left(\Pi\left(n(\mathbf{w} \langle \hat{\mathbf{H}}\mathbf{H}(\mathbf{x}) \rangle \langle \mathcal{X}_s^{i*} \rangle)\right)\right) - I^*(\mathbf{p}_i) \end{cases} \quad (7)$$

C. Efficient second-order minimization

Although the optimization problem can be solved using first-order methods [9], [10], [11], [12], we have chosen to use an efficient second-order minimization algorithm [14] [17]. Indeed, with little extra computation, we can make the most of the local quadratic convergence of the second-order optimization.

Consider the general minimization problem:

$$F(\mathbf{x}) = \frac{1}{2} \sum_{i=1}^n (f_i(\mathbf{x}))^2 = \frac{1}{2} \|\mathbf{f}(\mathbf{x})\|^2 \quad (8)$$

The necessary condition for finding a local minimum of the cost function is the existence of a stationary point \mathbf{x}_0 with:

$$[\nabla_{\mathbf{x}} F]_{\mathbf{x}=\mathbf{x}_0} = \mathbf{0} \quad (9)$$

where $\nabla_{\mathbf{x}}$ is the gradient operator with respect to the parameter \mathbf{x} . When equation (9) is non-linear, a closed-form solution is generally difficult to obtain.

A second-order Taylor series of \mathbf{f} about $\mathbf{x} = \mathbf{0}$ gives:

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{0}) + \mathbf{J}(\mathbf{0}) \mathbf{x} + \frac{1}{2} \mathbf{M}(\mathbf{0}, \mathbf{x}) \mathbf{x} + \mathcal{R}(\|\mathbf{x}\|^3) \quad (10)$$

where $\mathbf{J}(\mathbf{0}) = [\nabla_{\mathbf{x}} \mathbf{f}]_{\mathbf{x}=\mathbf{0}}$, $\mathbf{M}(\mathbf{z}, \mathbf{x}) = [\nabla_{\mathbf{x}} \mathbf{J}]_{\mathbf{x}=\mathbf{z}} \mathbf{x}$ and $\mathcal{R}(\|\mathbf{x}\|^3)$ is the third-order reminder. Similarly, we can write the Taylor series of the Jacobian \mathbf{J} about $\mathbf{x} = \mathbf{0}$:

$$\mathbf{J}(\mathbf{x}) = \mathbf{J}(\mathbf{0}) + \mathbf{M}(\mathbf{0}, \mathbf{x}) + \mathcal{R}(\|\mathbf{x}\|^2) \quad (11)$$

Plugging (11) in (10) leads to:

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{0}) + \frac{1}{2} (\mathbf{J}(\mathbf{0}) + \mathbf{J}(\mathbf{x})) \mathbf{x} + \mathcal{R}(\|\mathbf{x}\|^3) \quad (12)$$

Using equation (12), we obtain a second-order approximation of the cost function:

$$F(\mathbf{x}) = \frac{1}{2} \left\| \mathbf{f}(\mathbf{0}) + \frac{\mathbf{J}(\mathbf{0}) + \mathbf{J}(\mathbf{x})}{2} \mathbf{x} \right\|^2 + \mathcal{R}(\|\mathbf{x}\|^3) \quad (13)$$

Setting $\hat{\mathbf{f}}(\mathbf{x}) = \mathbf{f}(\mathbf{0}) + \frac{1}{2} (\mathbf{J}(\mathbf{0}) + \mathbf{J}(\mathbf{x})) \mathbf{x}$, the cost function is approximated by:

$$F(\mathbf{x}) \approx \frac{1}{2} \|\hat{\mathbf{f}}(\mathbf{x})\|^2$$

The derivative of the cost function is approximated by:

$$\nabla_{\mathbf{x}} F \approx (\nabla_{\mathbf{x}} \hat{\mathbf{f}})^T \hat{\mathbf{f}}$$

and using equation (9), we find the approximated local minimizer by solving the equation:

$$[\nabla_{\mathbf{x}} \hat{\mathbf{f}}]_{\mathbf{x}=\mathbf{x}_0}^T \left(\mathbf{f}(\mathbf{0}) + \frac{\mathbf{J}(\mathbf{0}) + \mathbf{J}(\mathbf{x}_0)}{2} \mathbf{x}_0 \right) = 0 \quad (14)$$

1) *Application to the tracking problem:* The Jacobians $\mathbf{J}(\mathbf{0})$ and $\mathbf{J}(\mathbf{x}_0)$, computed in the Appendix, can be written as the product of five Jacobians:

$$\mathbf{J}(\mathbf{0}) = \mathbf{J}_I \mathbf{J}_{\Pi} \mathbf{J}_n \mathbf{J}_w \mathbf{J}_H(\mathbf{0})$$

$$\mathbf{J}(\mathbf{x}_0) = \mathbf{J}_{I^*} \mathbf{J}_{\Pi} \mathbf{J}_n \mathbf{J}_w \mathbf{J}_{H^*}(\mathbf{x}_0)$$

Despite the Jacobian $\mathbf{J}(\mathbf{x}_0)$ generally depending on the unknown \mathbf{x}_0 , thanks to the left-invariance property of the vector fields on $\mathbb{SL}(3)$, the following identity can be proved:

$$\mathbf{J}_{H^*}(\mathbf{x}_0) \mathbf{x}_0 = \mathbf{J}_H(\mathbf{0}) \mathbf{x}_0$$

Thus, in the equation (14), we can use $\mathbf{J}_H(\mathbf{0}) \mathbf{x}_0$ instead of $\mathbf{J}_{H^*}(\mathbf{x}_0) \mathbf{x}_0$ where $\mathbf{J}_H(\mathbf{0})$ is a constant Jacobian matrix detailed in the Appendix. The update $\hat{\mathbf{x}}$ of the second-order minimization algorithm can be computed as follows:

$$\hat{\mathbf{x}} = - \left(\left(\frac{\mathbf{J}_I + \mathbf{J}_{I^*}}{2} \right) \mathbf{J}_{\Pi} \mathbf{J}_n \mathbf{J}_w \mathbf{J}_H(\mathbf{0}) \right)^+ \mathbf{f}(\mathbf{0})$$

The computational complexity is almost the same as for first-order algorithms especially as the reference Jacobian \mathbf{J}_{I^*} only needs to be calculated once.

Obtaining J_{I^*} and J_I , that are the Jacobians taken in the images, is a remarkable property. It indicates that we can take into account the non-linear properties of the sensor simply through the jacobian of the projection function \mathbf{J}_Π (this appears in the derivation of the jacobian in the Appendix).

IV. EXPERIMENTAL RESULTS

The algorithm was tested on real data obtained from a mobile robot. The central catadioptric camera was comprised of a S80 parabolic mirror from RemoteReality with a telecentric lens and perspective camera of resolution 1024×768 ¹.

The sequence was done over 120 images and a distance of about 2 m. Four reference templates were tracked (Fig. 3), they are numbered in counterclockwise order from 1 to 4 starting top left. Templates number 2 and number 3 were considered on the same plane (i.e. only one homography was estimated). For each homography, a translation \mathbf{t} up to a scale factor and a rotation \mathbf{R} can be extracted (the ambiguity was solved by using several frames). The scale was then fixed by measuring the distance of the camera to one of the planes. In Fig. 10 and Fig. 11 are shown the median of the estimated motions in dotted lines and the odometry in full lines.

The camera field of view was obstructed in the case of the template number 1 after the image 100 as we can see in Fig. 7, Fig. 8 and Fig. 9. The algorithm which uses a straight forward minimization was not able to find the correct homography, this does not appear in the motion estimation as a median is used.

Template number 4 was correctly tracked as we can see in Fig. 9, the complex motion depicted in the images (Fig. 4-8) are only due to the mirror geometry.

The estimated translation gave a maximum error of [2.7, 3.6, 7.3] cm on the [x,y,z] axis and an mean absolute error of [1, 1.3, 1.4] cm. The maximal error for the rotation was [1.6, 2.2, 1.0] deg with a mean absolute error of [0.8, 0.6, 0.3] deg. (The estimated angle obtained from the normals between templates 1 and 2 was of 87 deg and between templates 1 and 3 of 91 deg.)

V. CONCLUSION

We have presented in this article a homography-based approach to tracking for central catadioptric sensors. It has proved a precise method for motion estimation.

Future work will concentrate on making the present version robust to partial occlusion or illumination variation [10]. In order to improve motion estimation, euclidean constraints could also be integrated to link the motion of the different templates tracked.

APPENDIX: JACOBIANS COMPUTATION

For clarity, we will no longer indicate the index i for \mathbf{f}_i .

¹the camera was calibrated with the open-source toolbox available on <http://www-sop.inria.fr/icare/personnel/Christopher.Mei/Toolbox.html>

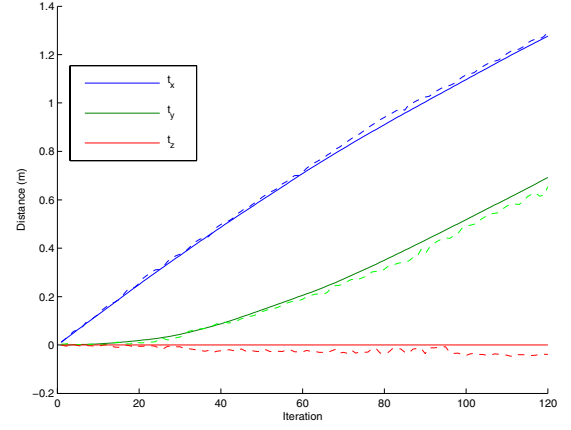


Fig. 10. Estimation of the translation of the robot

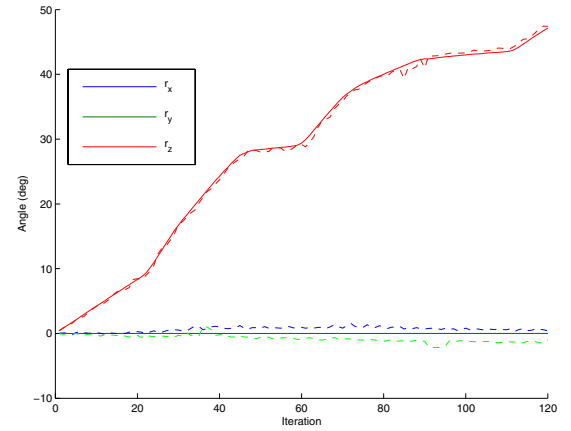


Fig. 11. Estimation of the rotation of the robot

Current Jacobian

We will write the current Jacobian as the product of five different Jacobians:

$$\begin{aligned} \mathbf{J}(\mathbf{0}) &= \left[\nabla_{\mathbf{x}} \left(I \left(\Pi \left(n(\mathbf{w} \langle \hat{\mathbf{H}} \mathbf{H}(\mathbf{x}) \rangle \langle \mathcal{X}_s^* \rangle) \right) \right) - I^*(\mathbf{p}) \right) \right]_{\mathbf{x}=\mathbf{0}} \\ &= \mathbf{J}_I \mathbf{J}_\Pi \mathbf{J}_n \mathbf{J}_w \mathbf{J}_H(\mathbf{0}) \end{aligned}$$

Noting that that :

$$\begin{aligned} \mathbf{w} \langle \hat{\mathbf{H}} \mathbf{H}(\mathbf{x}) \rangle \langle \mathcal{X}_s^* \rangle &= \mathbf{w} \langle \hat{\mathbf{H}} \rangle \langle \mathbf{w} \langle \mathbf{H}(\mathbf{x}) \rangle \langle \mathcal{X}_s^* \rangle \rangle \\ &= \mathbf{w} \langle \hat{\mathbf{H}} \rangle \langle \Pi^{-1}(\mathbf{q}) \rangle \end{aligned}$$

with $\mathbf{q} = \Pi(n(\mathbf{w} \langle \mathbf{H}(\mathbf{x}) \rangle \langle \mathcal{X}_s^* \rangle))$, the first Jacobian \mathbf{J}_I^* is:

$$\begin{aligned} \mathbf{J}_I^* &= \left[\nabla_{\mathbf{q}} \left(I \left(\Pi \left(n(\mathbf{w} \langle \hat{\mathbf{H}} \rangle \langle \Pi^{-1}(\mathbf{q}) \rangle) \right) \right) \right) \right]_{\mathbf{q}=\mathbf{t}} \\ \text{with : } \mathbf{t} &= \Pi(n(\mathbf{w} \langle \mathbf{H}(\mathbf{0}) \rangle \langle \mathcal{X}_s^* \rangle)) \end{aligned}$$

$\Pi(n(\mathbf{w} \langle \mathbf{H}(\mathbf{0}) \rangle \langle \mathcal{X}_s^* \rangle)) = \Pi(n(\mathbf{w} \langle \mathbf{I} \rangle \langle \mathcal{X}_s^* \rangle)) = \mathbf{p}$ so \mathbf{J}_I is the jacobian of the current image.

With $\Pi = k \circ h$:

$$\mathbf{J}_\Pi = \mathbf{J}_k \mathbf{J}_h = \frac{1}{Z_s - \xi} \begin{bmatrix} \gamma & \gamma s & -\frac{f_1(X_s + sY_s)}{Z_s - \xi} \\ 0 & \gamma r & -\frac{f_2 Y_s}{Z_s - \xi} \end{bmatrix}$$

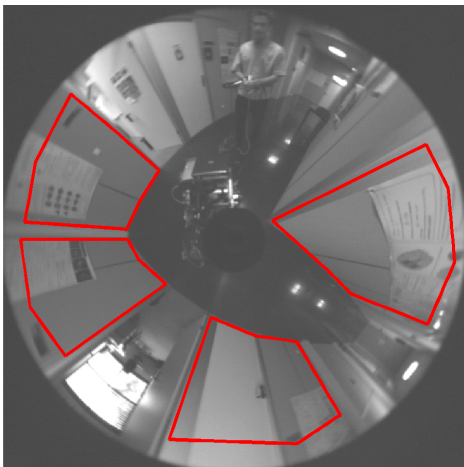


Fig. 3. Reference image

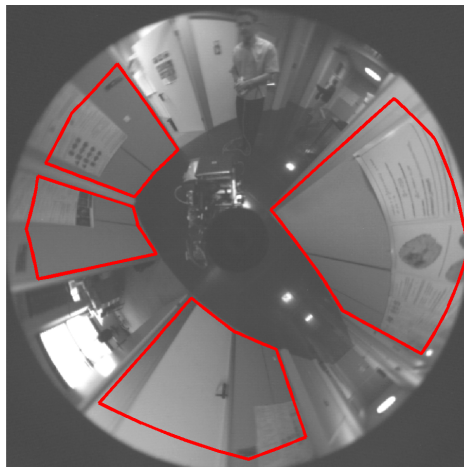


Fig. 4. Image 25

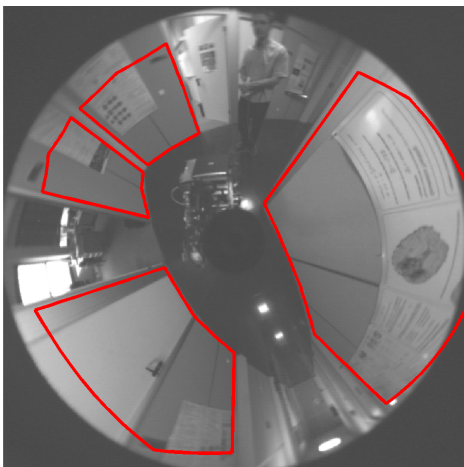


Fig. 5. Image 50

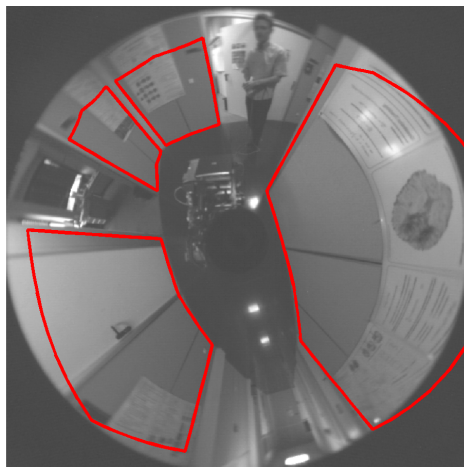


Fig. 6. Image 75

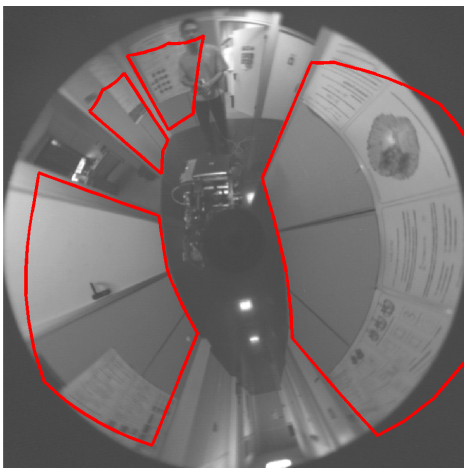


Fig. 7. Image 100

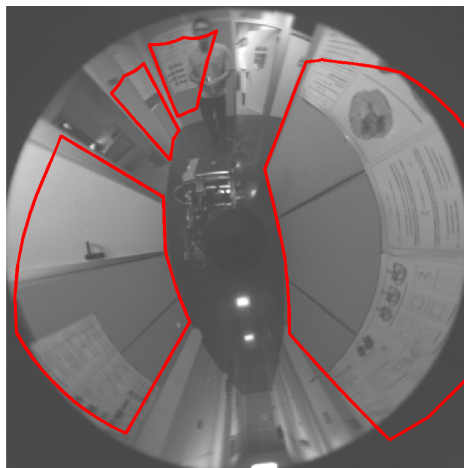


Fig. 8. Image 120

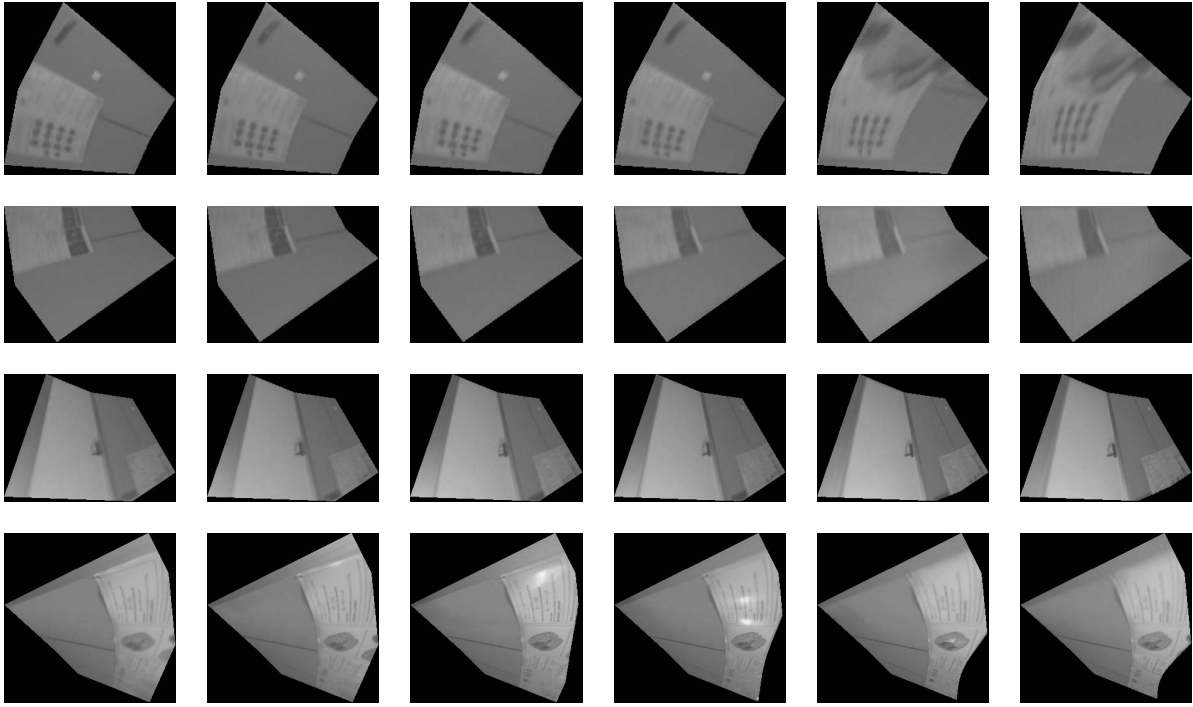


Fig. 9. Reprojection of the templates for iterations 0,25,50,75,100,120 in the reference image using the estimated homography

$$\mathbf{J}_w = [\nabla_{\mathbf{H}} \mathbf{w} \langle \cdot \rangle \langle \mathcal{X}_s^* \rangle]_{\mathbf{H}=\mathbf{H}(0)=\mathbf{I}} = \begin{bmatrix} \mathcal{X}_s^{*\top} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathcal{X}_s^{*\top} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathcal{X}_s^{*\top} \end{bmatrix}_{3 \times 9}$$

$$\mathbf{J}_{\mathbf{H}}(\mathbf{0}) = [\nabla_{\mathbf{x}} \mathbf{H}]_{\mathbf{x}=\mathbf{0}} = [\text{flat}(\mathbf{A}_1)^\top \text{flat}(\mathbf{A}_2)^\top \cdots \text{flat}(\mathbf{A}_8)^\top]_{9 \times 8}^\top$$

with : $\text{flat}(\mathbf{M}_{n \times m}) = [m_{11} \ m_{12} \ \cdots \ m_{1m} \ m_{21} \ m_{22} \ \cdots \ m_{nm}]^\top$

Reference Jacobian

The reference Jacobian $\mathbf{J}(\mathbf{x}_0)$ can be written as:

$$\mathbf{J}(\mathbf{x}_0) = \left[\nabla_{\mathbf{x}} \left(I \left(\Pi \left(n(\mathbf{w} \langle \hat{\mathbf{H}} \mathbf{H}(\mathbf{x}) \rangle \langle \mathcal{X}_s^* \rangle) \right) \right) - I^*(\mathbf{p}) \right) \right]_{\mathbf{x}=\mathbf{x}_0}$$

$$= \mathbf{J}_{I^*} \mathbf{J}_{\Pi} \mathbf{J}_n \mathbf{J}_w \mathbf{J}_{H^*}(\mathbf{x}_0)$$

Noting that :

$$\mathbf{w} \langle \hat{\mathbf{H}} \mathbf{H}(\mathbf{x}) \rangle \langle \mathcal{X}_s^* \rangle = \mathbf{w} \langle \bar{\mathbf{H}} \rangle \langle \mathbf{w} \langle \bar{\mathbf{H}}^{-1} \hat{\mathbf{H}} \mathbf{H}(\mathbf{x}) \rangle \langle \mathcal{X}_s^* \rangle \rangle$$

$$= \mathbf{w} \langle \bar{\mathbf{H}} \rangle \langle \Pi^{-1}(\mathbf{q}) \rangle$$

with $\mathbf{q} = \Pi(n(\mathbf{w} \langle \bar{\mathbf{H}}^{-1} \hat{\mathbf{H}} \mathbf{H}(\mathbf{x}) \rangle \langle \mathcal{X}_s^* \rangle))$, the Jacobian \mathbf{J}_I is:

$$\mathbf{J}_I = \left[\nabla_{\mathbf{q}} \left(I \left(\Pi \left(n(\mathbf{w} \langle \bar{\mathbf{H}} \rangle \langle \Pi^{-1}(\mathbf{q}) \rangle) \right) \right) \right) \right]_{\mathbf{q}=\mathbf{t}}$$

with : $\mathbf{t} = \Pi(n(\mathbf{w} \langle \bar{\mathbf{H}}^{-1} \hat{\mathbf{H}} \mathbf{H}(\mathbf{x}_0) \rangle \langle \mathcal{X}_s^* \rangle))$

$\Pi(n(\mathbf{w} \langle \bar{\mathbf{H}}^{-1} \hat{\mathbf{H}} \mathbf{H}(\mathbf{x}_0) \rangle \langle \mathcal{X}_s^* \rangle)) = \Pi(n(\mathbf{w} \langle \mathbf{I} \rangle \langle \mathcal{X}_s^* \rangle)) = \mathbf{p}$ so \mathbf{J}_I^* is the jacobian of the reference image.

$$\mathbf{J}_{H^*}(\mathbf{x}_0) = \left[\nabla_{\mathbf{x}} \bar{\mathbf{H}}^{-1} \hat{\mathbf{H}} \mathbf{H} \right]_{\mathbf{x}=\mathbf{x}_0}$$

Thanks to the Lie group parameterization, it can also be shown that the product $\mathbf{J}_{H^*}(\mathbf{x}_0) \mathbf{x}_0 = \mathbf{J}_H(\mathbf{0}) \mathbf{x}_0$.

REFERENCES

- [1] Y. Yagi, "Omnidirectional sensing and its applications," *IEICE Trans. on Information and Systems*, vol. E82-D, no. 3, pp. 568–579, 1999.
- [2] S. Baker and S. Nayar, "A theory of catadioptric image formation," in *ICCV*, 2001, pp. 422–427.
- [3] C. Geyer and K. Daniilidis, "A unifying theory for central panoramic systems and practical implications," in *ECCV*, 2000, pp. 445–461.
- [4] J. P. Barreto and H. Araujo, "Issues on the geometry of central catadioptric image formation," in *CVPR*, vol. 2, 2001, pp. 422–427.
- [5] T. Svoboda and P. T., "Epipolar geometry for central catadioptric cameras," *International Journal on Computer Vision*, 2002.
- [6] C. Geyer and K. Daniilidis, "Mirrors in motion: Epipolar geometry and motion estimation," *International Journal on Computer Vision*, 2003.
- [7] F. M. J. Barreto and R. Horaud, "Visual servoing/tracking using central catadioptric cameras," in *International Symposium on Experimental Robotics*, ser. Advanced Robotics Series, 2002.
- [8] H. Hadj-Abdelkader, Y. Mezouar, N. Andreff, and P. Martinet, "2 1/2 d visual servoing with central catadioptric cameras," in *IROS*, 2005.
- [9] B. Lucas and T. Kanade, "An iterative image registration technique with application to stereo vision," in *International Joint Conference on Artificial Intelligence*, 1981, pp. 674–679.
- [10] G. D. Hager and P. N. Belhumeur, "Efficient region tracking with parametric models of geometry and illumination," *Trans. on Pattern Analysis and Machine Intelligence*, 1998.
- [11] H. Y. Shum and R. Szeliski, "Construction of panoramic image mosaics with global and local alignment," *International Journal on Computer Vision*, vol. 16, no. 1, pp. 63–84, 2000.
- [12] S. Baker and I. Matthews, "Equivalence and efficiency of image alignment algorithms," in *CVPR*, 2001, pp. 1090–1097.
- [13] J. M. Buenaposada and L. Baumela, "Real-time tracking and estimation of planar pose," in *ICPR*, 2002, pp. 697–700.
- [14] E. Malis, "Improving vision-based control using efficient second-order minimization techniques," in *ICRA*, 2004.
- [15] H. Jin, P. Favaro, and S. Soatto, "A semi-direct approach to structure from motion," *The Visual Computer*, vol. 19, no. 6, pp. 377–394, 2003.
- [16] O. Faugeras and F. Lustman, "Motion and structure from motion in a piecewise planar environment," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 2, no. 3, pp. 485–508, 1988.
- [17] S. Benhimane and E. Malis, "Real-time image-based tracking of planes using efficient second-order minimization," in *IROS*, 2004.