The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

# Visual Tracking of Planes with an Uncalibrated Central Catadioptric Camera

A. Salazar-Garibay, E. Malis and C. Mei

*Abstract*— **This paper addresses the problem of tracking a planar region of the scene using an uncalibrated omnidirectional camera. Omnidirectional cameras are a popular choice of visual sensors in robotics because the large field of view is well adapted to motion estimation and obstacle avoidance. The novelty of this work resides in simplifying the calibration phase by providing a direct approach to tracking without any prior knowledge of the camera, lens or mirror parameters. We deal with a nonlinear optimization problem that can be solved for small displacements between two images like those acquired at video rate by a camera mounted on a robot. In order to assess the performance of the proposed method, we perform experiments with synthetic and real data.**

## I. INTRODUCTION

Omnidirectional cameras are important in areas where large visual field coverage is needed, such as motion estimation and obstacle avoidance [13], [3]. However their practical use is often burdened by the calibration phase that can be time consuming and require an experienced user. The motivation for this work is thus to simplify this phase by providing a direct approach to tracking a planar region without any prior knowledge of the camera, lens or mirror parameters.

Visual tracking methods can be classified in two main groups: *feature-based* methods and *direct* approaches. In the first case, local features such as points, line segments, edges or contours are tracked across the sequence. By using adequate descriptors and extractors (such as SIFT) combined with robust computation using for example RANSAC, wide baseline uncalibrated omnidirectional structure from motion is possible. Successful approaches have been demonstrated by extending the concept of fundamental matrix leading to poly-eigenvalue problems that can be solved efficiently [5], [12]. These approaches however require specialized image processing tools due to the anisotropic resolution of paracatadioptric cameras. The second group of methods minimises a cost function based on the image data directly. The approach studied in this article uses gradient descent to minimise the sum of squared differences of the image intensities as in [8], [7], [4], [1], [10]. It assumes small inter-frame motion such as what is typically produced by a 30Hz camera mounted on a mobile robot. The advantages of this technique compared to feature-based methods are the sub-pixel accuracy and the high frame-rate. The disadvantages come from the assumption on small inter-frame motion that makes it unable to cope with rapid motion. In [10], the authors proposed a direct method for tracking piecewise planar objects with a central catadioptric camera. They extended the standard notion of homography to omnidirectional cameras using the unified projection model on the sphere. One of the limitations of this work was the need for a precisely calibrated omnidirectional sensor.

The aim of this article is to show how this assumption can be lifted paving the way for a direct approach to uncalibrated structure from motion. The proposed approach minimises the sum of squared differences (SSD) between a region in a reference image and a warped region of the current image. We deal with a nonlinear optimization problem that can be solved for small displacements between two images acquired by a camera mounted on a robot. In order to solve the least-squares optimization problem we apply the efficient second-order minimization method (ESM) [9]. Experimental results show that, unlike previous work, our method is able to track planar objects with an uncalibrated catadioptric camera and thus can be helpful in robotic applications where camera calibration is impossible or hard to obtain.

The paper is organized as follows. In Section II we describe the projection model. The main contribution in this paper, uncalibrated visual tracking, is described in Section III. The experimental results are shown in Section IV. Finally, Section V concludes the paper and presents ideas for future work.

## II. THEORETICAL BACKGROUND

### A. Projection Model

This section describes the projection model and defines the calibration parameters. We followed the model proposed in [11] that is a slightly modified version of the projection model of Geyer [6] and Barreto [2]. We assume the camera and mirror to be a single imaging device and use a generalized focal length that is a product between the camera focal length and a parameter that defines the shape of the mirror. The projection of a 3D point into the image can be modeled as follows.

Let $\mathbf{m}$ be a 3D point having Cartesian coordinates $\mathbf{m} = (X,\ Y,\ Z)^\top$ in the camera frame $C_m$. The point is projected onto the point $\mathbf{s} = (X_s,\ Y_s,\ Z_s)^\top$ on the unit sphere $S$ centered at the origin: $\mathbf{s} = \frac{\mathbf{m}}{\|\mathbf{m}\|}$. The point on the sphere is then projected from a point at a distance $\xi$ from the origin of the sphere. $\xi$ depends on the mirror parameters as described

A. Salazar-Garibay is with INRIA Sophia-Antipolis, France. Adan.Salazar@sophia.inria.fr

E. Malis is with INRIA Sophia-Antipolis, France. Ezio.Malis@sophia.inria.fr

C. Mei is with the University of Oxford, United Kingdom. christopher.mei@eng.ox.ac.uk

This work is partially supported by the grants CONACYT and SEP, México.

978-1-4244-3804-4/09/$25.00 ©2009 IEEE

in Table I. Let $\mathbf{q} = (q_u,\ q_v)^\top$ be the new point after this projection: $\mathbf{q} = \mathbf{h}(\xi, \mathbf{s}) = \left(\frac{X_s}{Z_s-\xi},\ \frac{Y_s}{Z_s-\xi}\right)^\top$

TABLE I

MIRROR PARAMETER

| $\xi$ | **Mirror type** |
|---|---|
| 1 | Parabolic |
| >0 and <1 | Hyperbolic, elliptic, conical or spherical |
| 0 | Planar |
| >1 | Fish eye |
| <0 | No mirror |

Finally, we obtain the coordinates $(u, v)$ of an image point $\mathbf{p} = \mathbf{k}(\boldsymbol{\gamma}, \mathbf{q}) = \begin{bmatrix} k_u & 0 & u_0 \\ 0 & k_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{q}$.

$k_u$ and $k_v$ represent the generalized focal length and $(u_0,\ v_0)$ is the principal point (in pixels). $\boldsymbol{\gamma}$ contains the camera intrinsic parameters.

From a point $\mathbf{p}$ measured in the image, it is possible to lift it to the unit sphere. The first step is to apply the inverse projection induced by $\mathbf{k}$. We then obtain a point on the normalized plane $\mathbf{q} = \mathbf{k}^{-1}(\boldsymbol{\gamma}, \mathbf{p})$. The second step projects this on the unit sphere using the inverse function proposed by Barreto [2]: $\mathbf{s} = \mathbf{h}^{-1}(\xi, \mathbf{q}) = (\alpha\ q_u,\ \alpha\ q_v,\ \alpha + \xi)^\top$ with $\alpha = -\frac{\xi+\sqrt{1+(1-\xi^2)(q_u^2+q_v^2)}}{q_u^2+q_v^2+1}$

### B. Motion Model

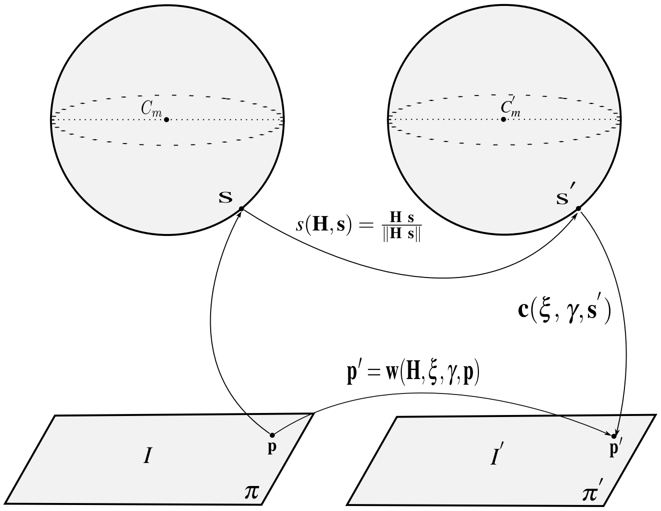
If we suppose that the camera observes a planar object in the scene, the displacement of a point on the sphere can be represented by a homography. The homography contains the $3 \times 3$ rotation matrix $\mathbf{R}$ of the camera and its translation vector $\mathbf{t}$. Figure 1 illustrates the transformation induced by a planar homography using the spherical projection model. Two planar points are related by a homography $\mathbf{H}$ by $\mathbf{X}' = \mathbf{HX}$, so the projection of points $\mathbf{s}$ and $\mathbf{s}'$, belonging to a planar region of the scene, on the sphere are related by $\rho'\mathbf{s}' = \rho\mathbf{H}\mathbf{s}$. 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 $\mathbf{n}^*$ (a unit vector) and the distance $d^*$ of the plane to the origin of the reference frame.

### C. Warping

Warping will be for us a function that allows to find the coordinates of reference image points in the current image (See Figure 1). We will denote by $\mathbf{w}$ the warping function which depends on the homography and the sensor parameters:

$$\begin{array}{rll} \mathbf{w}\colon \mathbb{SL}(3)\times\mathbb{R}\times\mathbb{R}^6\times\mathbb{R}^2 & \rightarrow & \mathbb{P}^2 \\ (\mathbf{H}, \xi, \boldsymbol{\gamma}, \mathbf{p}) & \rightarrow & \mathbf{p}^{'} = \mathbf{w}\left(\mathbf{H}, \xi, \boldsymbol{\gamma}, \mathbf{p}\right) \end{array}$$

The steps of warping function include basically three transformations: 1) The transformation between the image plane and the unit sphere, 2) The transformation between spheres and 3) The transformation between the unit sphere and the image plane. Figure 1 depicts the three transformations.



$C_m$
$C_m^{'}$
$\mathbf{s}$
$\mathbf{s}^{'}$
$s(\mathbf{H},\mathbf{s}) = \frac{\mathbf{H}\ \mathbf{s}}{\|\mathbf{H}\ \mathbf{s}\|}$
$\mathbf{c}(\xi,\ \gamma,\mathbf{s}^{'})$
$\mathbf{p}^{'} = \mathbf{w}(\mathbf{H},\xi,\gamma,\mathbf{p})$
$I$
$\mathbf{p}$
$\pi$
$I^{'}$
$\mathbf{p}^{'}$
$\pi^{'}$

Fig. 1. **Motion model**. Transformation induced by a planar homography using the spherical projection model. The points $\mathbf{s}$ and $\mathbf{s}'$ are related by $\rho'\mathbf{s}' = \rho\mathbf{H}\mathbf{s}$.

Let $\mathbf{c}(\xi, \boldsymbol{\gamma}, \mathbf{s}') = \mathbf{k}(\boldsymbol{\gamma}, \mathbf{h}(\xi, \mathbf{s}'))$ be the transformation between the sphere and the image plane:

$$\begin{array}{rcl} \mathbf{c}: \quad \mathbb{R}\times\mathbb{R}^4\times\mathbb{S}^2 & \longrightarrow & \mathbb{P}^2 \\ (\xi, \boldsymbol{\gamma}, \mathbf{s}') & \longrightarrow & \mathbf{p} = \mathbf{c}(\xi, \boldsymbol{\gamma}, \mathbf{s}') \end{array}$$

This transformation uses $\mathbf{h}(\xi, \mathbf{s})$ and $\mathbf{k}(\boldsymbol{\gamma}, \mathbf{q})$ to project a point from the unit sphere to the image. The inverse of this transformation applies the inverse projection induced by $\mathbf{k}^{-1}(\boldsymbol{\gamma}, \mathbf{p})$ and $\mathbf{h}^{-1}(\xi, \mathbf{q})$ to lift a point from the image to the unit sphere.

For the transformation between spheres let $\boldsymbol{\Psi}(\mathbf{H}, \mathbf{s}) = \frac{\mathbf{Hs}}{\|\mathbf{Hs}\|}$ be the function that transforms the points $\mathbf{s}$ and $\mathbf{s}'$ between the spheres:

$$\begin{array}{rcl} \boldsymbol{\Psi}: \quad \mathbb{SL}(3)\times\mathbb{S}^2 & \longrightarrow & \mathbb{S}^2 \\ (\mathbf{H}, \mathbf{s}) & \longrightarrow & \mathbf{s}' = \boldsymbol{\Psi}(\mathbf{H}, \mathbf{s}) \end{array}$$

If theses expressions are combining, the warping function can be written as:

$$\mathbf{w}(\mathbf{H}, \xi, \boldsymbol{\gamma}, \mathbf{p}) = \mathbf{c}(\xi, \boldsymbol{\gamma}, \boldsymbol{\Psi}(\mathbf{H}, \mathbf{c}^{-1}(\xi, \boldsymbol{\gamma}, \mathbf{p}))) \qquad (1)$$

This warping expression will be higly usefull in the rest of paper.

## III. UNCALIBRATED VISUAL TRACKING

The tracking problem will essentially be considered as an image registration problem which will be related directly to the grey-level brightness measurements in the catadioptric images via the non-linear model presented in section II which accounts for the model formation of the image. Since our final objective is to track a plane in catadioptric images an image reference of that plane is considered.

Let $I$ be the reference image. We will call *reference template*, a region of size $\mathcal{R}$ (rows × columns) of $I$ corresponding to the projection of a 3D planar region of the scene. To track the reference template in the current image $I^{'}$, we look for a set of parameters including the mirror parameter $\xi$, the camera intrinsic parameters $\boldsymbol{\gamma}$ and the transformation $\mathbf{H}$ such that current image will be aligned with the reference template.

$$I'\left(\mathbf{w}\left(\mathbf{H}, \xi, \boldsymbol{\gamma}, \mathbf{p}\right)\right) = I(\mathbf{p}) \qquad (2)$$

These parameters needn't be unique. For example, in the perspective case, two views are not sufficient to estimate the camera intrinsic parameters. Our aim in this article is not to recover the true parameters but to align the image regions. Once we have an approximation $\widehat{\mathbf{H}}$ of the transformation $\mathbf{H}$ and an approximation $\widehat{\xi}$ and $\widehat{\boldsymbol{\gamma}}$ of the intrinsic parameters $\xi$ and $\boldsymbol{\gamma}$, the problem is to find the incremental transformation of $\mathbf{H}$, $\xi$, and $\boldsymbol{\gamma}$, that minimize the sum of square differences over all the pixels of the cost function:

$$\frac{1}{2}\sum_{\mathbf{p}\in\mathcal{R}} \|I^{'}\left(\mathbf{w}\left(\widehat{\mathbf{H}}\mathbf{H}, \widehat{\xi}+\xi, \widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma}, \mathbf{p}_i\right)\right) - I(\mathbf{p}_i)\|^2 \qquad (3)$$

The homography and intrinsic parameters of the imaging device are then updated as follows:

$$\begin{array}{ccc} \widehat{\mathbf{H}} & \leftarrow & \widehat{\mathbf{H}}\mathbf{H} \\ \widehat{\xi} & \leftarrow & \widehat{\xi}+\xi \\ \widehat{\boldsymbol{\gamma}} & \leftarrow & \widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma} \end{array} \qquad (4)$$

Similarly to [10] the incremental homography $\mathbf{H} \in \mathbb{SL}(3)$ and the intrinsic parameters updated are parametrized with local coordinates of the Lie algebra $\mathfrak{sl}(3)$. However, let us remark that the optimization problem is much more challenging than the problem solved in [10] where only the homography was computed because they used a calibrated camera (known intrinsic parameters).

In this work we do not attempt to estimate the motion and structure directly but estimate the relative motion as a homography. Although this approach does not use the minimal amount of parameters and could lead to a less stable estimate, it circumvents the issue of choosing the correct homography decomposition and deciding when reliable 3D tracking can occur.

### *A. Application of the Efficient Second-order Method (ESM)*

The aim now is to minimise the objective criterion defined previously in equation (3) in an accurate and robust manner. As this is a non-linear function of the unknown parameters, an iterative procedure is employed. Let $\mathbf{x} = (\mathbf{z}, \xi, \boldsymbol{\gamma})$ be the state vector. $\mathbf{z}$ contains the homography parameters. The objective function is minimized by: $\nabla_{\mathbf{x}}\mathbf{f}_i(\mathbf{x})|_{\mathbf{x}=\widetilde{\mathbf{x}}} = \mathbf{0}$, where $\nabla_{\mathbf{x}}$ is the gradient operator with respect to the unknown parameters and there exists a stationary point $\mathbf{x} = \widetilde{\mathbf{x}}$ which is the global minimum of the cost function.
Since both the reference image and current image are available it is possible to use the efficient second-order approximation method (ESM) [9] to solve the optimization problem. In this case the current and reference jacobians are:

$$\mathbf{J}(\mathbf{0}) = \mathbf{J}_{I'}\mathbf{J}_{\mathbf{w}}\left[\begin{array}{ccc} \mathbf{J}_{\mathbf{w}_\mathbf{H}}\mathbf{J}_{\mathbf{H}}(\mathbf{0}) & \mathbf{J}_{\xi}(\mathbf{0}) & \mathbf{J}_{\boldsymbol{\gamma}}(\mathbf{0}) \end{array}\right] \qquad (5)$$

$$\mathbf{J}(\tilde{\mathbf{x}}) = \mathbf{J}_I\mathbf{J}_{\mathbf{w}}\left[\begin{array}{ccc} \mathbf{J}_{\mathbf{w}_\mathbf{H}}\mathbf{J}_{\mathbf{H}^{-1}\widehat{\mathbf{H}}}(\tilde{\mathbf{x}}) & \mathbf{J}_{\xi}(\tilde{\mathbf{x}}) & \mathbf{J}_{\boldsymbol{\gamma}}(\tilde{\mathbf{x}}) \end{array}\right] \qquad (6)$$

Because $\mathbf{J}(\tilde{\mathbf{x}})$ depends explicity on the unknown optimal increment $\tilde{\mathbf{x}}$ we use the left invariance property $\mathbf{J}(\tilde{\mathbf{x}})\tilde{\mathbf{x}} = \mathbf{J}(0)\tilde{\mathbf{x}}$ in order to avoid the computation of $\mathbf{J}_{\mathbf{H}^{-1}\widehat{\mathbf{H}}}(\tilde{\mathbf{x}})$, $\mathbf{J}_{\widehat{\xi}}(\tilde{\mathbf{x}})$ and $\mathbf{J}_{\widehat{\boldsymbol{\gamma}}}(\tilde{\mathbf{x}})$ by assuming $\mathbf{H} \approx \widehat{\mathbf{H}}$, $\xi \approx \widehat{\xi}$ and $\boldsymbol{\gamma} \approx \widehat{\boldsymbol{\gamma}}$. Therefore, the update $\tilde{\mathbf{x}}$ of the solution can then be computed as follows:

$$\widetilde{\mathbf{x}} = \left(\left(\frac{\mathbf{J}_I + \mathbf{J}_{I'}}{2}\right)\mathbf{J}_{\mathbf{w}}\left[\mathbf{J}_{\mathbf{w}_\mathbf{H}}\mathbf{J}_{\mathbf{H}}(\mathbf{0}) \;\; \mathbf{J}_{\xi}(\mathbf{0}) \;\; \mathbf{J}_{\boldsymbol{\gamma}}(\mathbf{0})\right]\right)^{+}\mathbf{f}(\mathbf{0}) \qquad (7)$$

where '+' indicates the matrix pseudo-inverse.

$\mathbf{J}_{I'}$ represents the current image gradient $\nabla_{I'}$ evaluated on a point $\mathbf{p}$ of the reference image $I$. $\mathbf{J}_{\mathbf{w}}$ represents the variation from a point $\mathbf{p}'$ in the current image $I'$ with respect to a point $\mathbf{p}$ in the reference image $I$. $\mathbf{J}_{\mathbf{w}_\mathbf{H}}\mathbf{J}_{\mathbf{H}}$ represents the variation from a point $\mathbf{p}^{'}$ in the current image $I'$ with respect to the homography parameters. $\mathbf{J}_{\xi}$ and $\mathbf{J}_{\boldsymbol{\gamma}}$ represent the variation from a point $\mathbf{p}'$ in the current image $I'$ with respect to the mirror parameter $\xi$ and the intrinsic parameters $\boldsymbol{\gamma}$ respectively. $\mathbf{J}_I$ is the Jacobian of the image reference and therefore only needs to be calculated once. The rest of Jacobians are recomputed at each iteration.

## IV. Results

In order to assess the performance of the proposed method we performed experiments with synthetic and real data. The synthetic image sequence is composed of 100 images. To create this sequence we transformed a real parabolic image. The synthetic images were created from this image assuming constant intrinsic parameters such as: a catadioptric camera with a parabolic mirror ($\xi = 1$), a generalized focal length $k_u = -250$, $k_v = -250$ and an image center $(u_0, v_0) = (512, 384)$. The homography matrices are different for each image. The real image sequence is composed of 120 images of size is $1024 \times 768$ combining a camera with a parabolic mirror. In both experiments, we compared our results with the visual tracking algorithm proposed in [10] where the authors assumed known intrinsic camera parameters.

### *A. Synthetic data*

In the first experiment with synthetic data we considered known intrinsic parameters. The assumed constant intrinsic parameters for both methods were $\xi = 1$, $k_u = -250$, $k_v = -250$ , $u_0 = 512$ , $v_0 = 384$. The initial guess for the homography parameters was given by the identity $3 \times 3$ matrix. Figure 2 shows the reprojection error (norm) using the visual tracking algorithm proposed in [10] (left) and for the method proposed in this paper (right). For all the sequence, the reprojection errors are almost the same.
The second experiment with synthetic data achieves with the aim of this paper. Therefore, we considered unknown intrinsic parameters to start the minimization method so,

we gave an initial guess of $\xi = 0.7$, $k_u = -100$, $k_v = -100$, $u_0 = 506$ and $v_0 = 375$. The initial guess for the homography parameters was given by the identity $3 \times 3$ matrix. Figure 3 shows 6 images of the test sequence with the tracked image region marked in red for the method in [10]. Figure 4 shows 6 images of the test sequence with the tracked image region marked in red for the proposed method in this paper. Figure 5 shows the reprojection error (norm) using the visual tracking algorithm proposed in [10] (left) and for the method proposed in this paper (right). As we can see, the proposed method in [10] is not able to track accurately the expected path along the sequence. On the other hand, even if the camera is not calibrated, the proposed algorithm is capable to track accurately the chosen plane along the sequence.

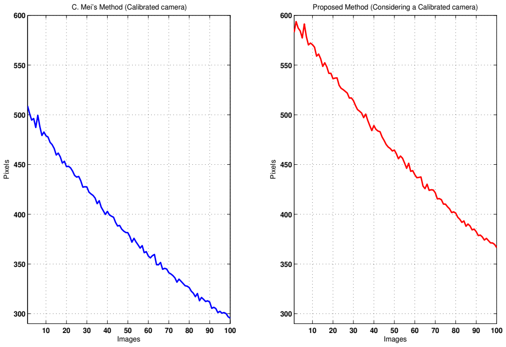


Fig. 2. **Reprojection error**. Reprojection error (norm) using the visual tracking algorithm proposed in [10] (left) and for the method proposed in this paper (right). The method propose in [10] computes only the homography parameters. We compute the homography parameters and the intrinsic parameters. For all the sequence, the reprojection errors are almost the same.

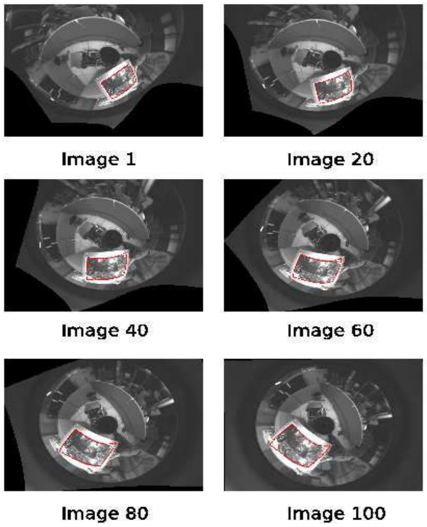


Fig. 3. **Visual tracking**. Plane tracked using the proposed method in [10]. It only computes the homography parameters. We supposed unknown intrinsic parameters.

## B. Real data

For this experiment we only considered the case with unknown intrinsic parameters. To start the minimization method we gave an initial guess of $\xi = 0.7$, $k_u = -100$, $k_v = -100$, $u_0 = 512$ and $v_0 = 384$. The initial guess for the homography parameters was given by the identity $3 \times 3$ matrix. Figure 6 shows 6 images of the real sequence with the tracked image region marked in red for the method in [10]. Figure 7 shows 6 images of the real sequence with the tracked image region marked in red for the proposed method in this paper. Figure 8 shows the reprojection error (norm) using the visual tracking algorithm proposed in [10] (left) and for the method proposed in this paper (right). As we can see, the proposed method in [10] is robust to track at least 50 images with unknown intrinsic parameters

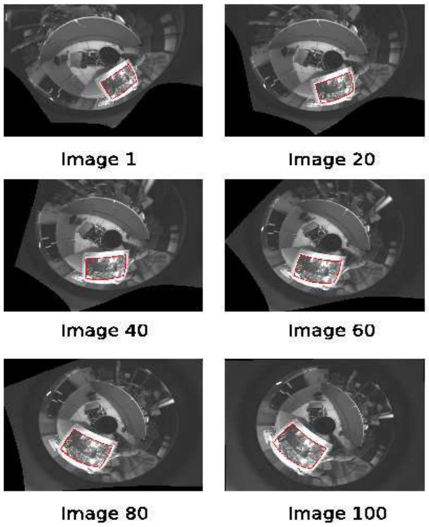


Fig. 4. **Visual tracking**. Plane tracked using the proposed visual tracking algorithm. We compute the homography parameters and the intrinsic parameters $\xi$, $k_u$, $k_v$, $u_0$ and $v_0$.

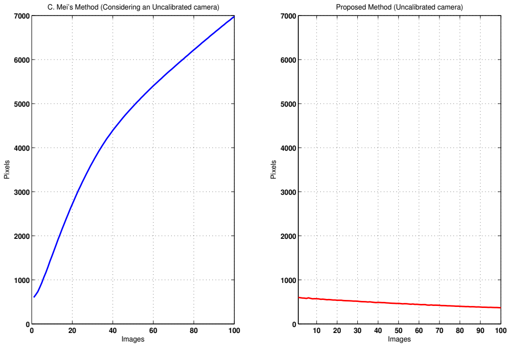


Fig. 5. **Reprojection error**. Reprojection error (norm) using the visual tracking algorithm proposed in [10] (left) and for the method proposed in this paper (right). The propoposed method in this paper computes the homography parameters and the intrinsic parameters $\xi$, $k_u$, $k_v$, $u_0$ and $v_0$. The proposed method in [10] does not compute the intrinsic parameters. For all the sequence, the reprojection error is different.

computing only the homography parameters. However, it start to loose the expected path after 60 images. That means that the homography matrix is not enough to minimise the reprojection error between the reference image and the current image while the displacement is increasing. On other hand, even if the camera is not calibrated, the proposed algorithm is capable to track accurately the choosen plane along the sequence.

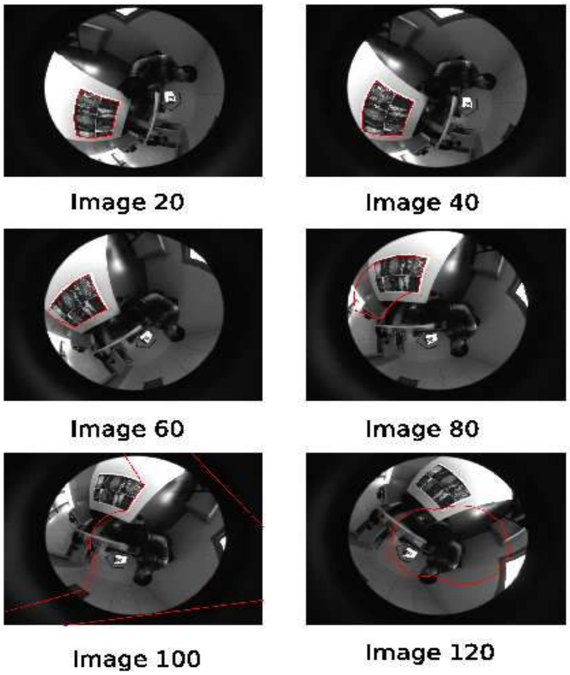


Image 20 Image 40
Image 60 Image 80
Image 100 Image 120

Fig. 6. **Visual tracking**. Plane tracked using the proposed method in [10]. It only computes the homography parameters. We supposed unknown intrinsic parameters. We supposed unknown intrinsic parameters.

## V. Conclusion

In this paper, we have shown how to efficiently track a plane in an omnidirectional image without requiring the prior calibration of the sensor. On the other hand, a set of required parameters are estimated on-line for each new image to align the current image with a reference template. The approach is very interesting because the estimated parameters are integrated into a single global warping function and we developed the Jacobian matrix of this warping function in easy modular parts. Furthermore, the efficient second order minimisation technique was applied in order to allow us minimisation of a highly redundant non-linear function in a precise manner. Avoiding the awkward calibration steps should facilitate the adoption of omnidirectional sensors in robotics. Future work will focalise on the self-calibration of the sensor on-line by using several of the tracked views. This should also enable to fix the values being estimated, providing a faster and more robust algorithm.

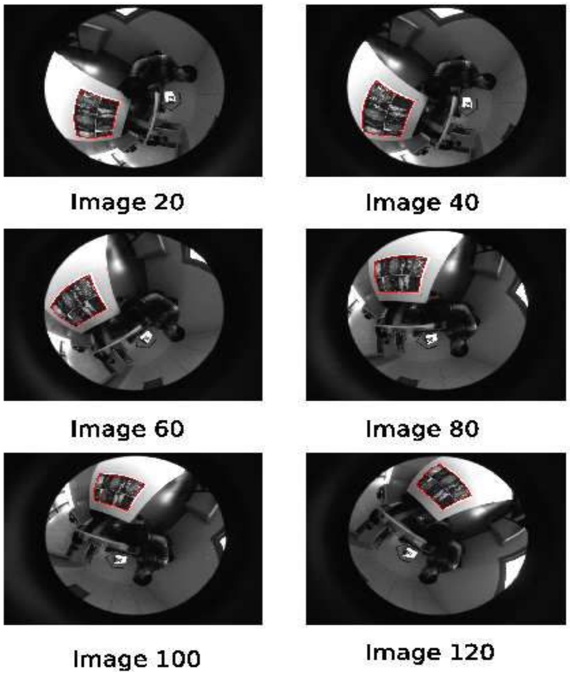


Image 20 Image 40
Image 60 Image 80
Image 100 Image 120

Fig. 7. **Visual tracking**. Plane tracked using the proposed visual tracking algorithm. We compute the homography parameters and the intrinsic parameters $\xi$, $k_u$, $k_v$, $u_0$ and $v_0$.

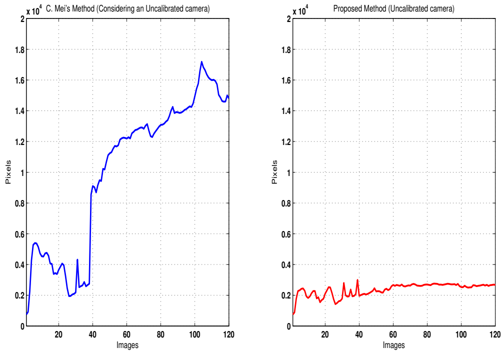


Fig. 8. **Reprojection error**. Reprojection error (norm) using the visual tracking algorithm proposed in [10] (left) and for the method proposed in this paper (right). The propoposed method in this paper computes the homography parameters and the intrinsic parameters $\xi$, $k_u$, $k_v$, $u_0$ and $v_0$. The proposed method in [10] does not compute the intrinsic parameters. For all the sequence, the reprojection error is different.

# References

[1] S. Baker and I. Matthews. Equivalence and efficiency of image alignment algorithms. In *CVPR*, pages 1090–1097, 2001.

[2] J. Barreto and H. Araujo. Issues on the geometry of central catadioptric image formation. In *CVPR*, volume 2, pages 422–427, 2001.

[3] R. Benosman and S. B. Kang. A brief historical perspective on panorama. In *Panoramic Vision*, pages 5–20. Apr 2001.

[4] J. M. Buenaposada and L. Baumela. Real-time tracking and estimation of planar pose. In *ICPR*, pages 697–700, 2002.

[5] A. W. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. In *CVPR*, 2001.

[6] C. Geyer and K. Daniilidis. A unifying theory for central panoramic systems and practical applications. In *European Conference on Computer Vision*, pages 445–461, 2000.

[7] G. D. Hager and P. N. Belhumeur. Efficient region tracking with parametric models of geometry and illumination. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(10):1025–1039, 1998.

[8] B. D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. *International Joint Conference on Artificial Intelligence*, pages 674–679, 1981.

[9] E. Malis. Improving vision-based control using efficient second-order minimization techniques. In *IEEE International Conference on Robotics and Automation*, 2004.

[10] C. Mei, S. Benhimane, E. Malis, and P. Rives. Efficient homography-based tracking and 3-d reconstruction for single-viewpoint sensors. *IEEE Transactions on Robotics*, 24(6):1352–1364, Dec. 2008.

[11] C. Mei and P. Rives. Single view point omnidirectional camera calibration from planar grids. In *IEEE International Conference on Robotics and Automation*, April 2007.

[12] B. Micusik and T. Pajdla. Structure from motion with wide circular field of view cameras. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(7), 2006.

[13] Y. Yagi. Omnidirectional sensing and its applications. *IEICE Trans, on Information and Systems*, E82-D(3):568–579, 1999.