Direct Approach to the Self-Calibration of Omnidirectional Cameras

A. Salazar-Garibay *
INRIA Sophia Antipolis
2004, route des Lucioles BP-93
06902, Sophia Antipolis, France
Adan.Salazar@sophia.inria.fr

E. Malis
INRIA Sophia Antipolis
2004, route des Lucioles BP-93
06902, Sophia Antipolis, France
Ezio.Malis@sophia.inria.fr

Abstract

Omnidirectional cameras are a popular choice of visual sensors in robotics because their large field of view is well suited for motion estimation and obstacle avoidance. However their practical use is often burdened by the calibration phase that can be time consuming and require an experienced user. The contribution of this work is a simplification of calibration phase by providing a direct approach to the calibration of this kind of sensor. This approach tracks a planar region in the scene to find the camera's intrinsic parameters. The only requirement for the user is to select a plane in the first image of the sequence. The algorithm needs no prior knowledge about extrinsic, camera, lens or mirror parameters. In order to assess the performance of the proposed method, we perform experiments with synthetic and real data.

1. Introduction

Since their introduction to the computer vision community, omnidirectional cameras have been utilized in many application areas where large visual field coverage is needed such as motion estimation and obstacle avoidance [15, 4]. However their practical use is often burdened by the calibration phase that can be time consuming and require an experienced user. Accurate calibration of a vision system is necessary for any computer vision task requiring extracting metric information of the environment from 2D images. Whereas existing works can provide interesting results, they suffer from several practical limitations (manual extraction, inaccurate conic fitting, calibration pattern, camera motion, execution time, etc ...).

Many calibration methods for omnidirectional cameras have been presented in the past few years. Previous works on omnidirectional camera calibration can be classified into two different categories. The first one includes methods which exploit prior knowledge about the scene, such as the presence of calibration patterns [6, 2] or plum lines [8] . The second group covers techniques that do not use this knowledge. This includes calibration methods from pure rotation [2] or planar motion of the camera [9] and self-calibration procedures, which are performed from point correspondences and epipolar constraint through minimizing an objective function [10, 14].

All mentioned techniques differentiate themselves mainly by: the type of mirror taken into account (e.g., hyperbolic and parabolic) and the projection model (skewness, aligment, errors, ...). In [8, 10], the authors treat the case of a parabolic mirror. In [8] it is shown that vanishing points lie on a conic section which encodes the entire calibration information. Thus, projections of two sets of parallel lines suffice for intrinsic calibration. Howewer, this property does not apply to non-parabolic mirros. Therefore, the proposed technique cannot be easily generalized to other kinds of sensors. The methods described in [10, 14] fall in the self-calibration category. These methods require no calibration pattern, nor a priori knowledge about the scene. The only assumption is the capability to automatically find point correspondences in a set of panoramic images of the same scene. Then, calibration is directly performed by epipolar geometry by minimizing an objective function. In [10], this is done by employing a parabolic mirror. However, besides focusing on particular types, the mentioned self-calibration techniques may suffer in case of tracking difficulties and of a small number of reatures points [5].

Motivated by this observation, we propose a calibration that stays valid for all central catadioptric systems. We use the unified model of Barreto-Geyer [3, 7] to achieve this. The proposed self-calibration method is based on previous work [1]. Here, a planar region in the scene was tracked using

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

2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops
978-1-4244-4441-0/09/$25.00 ©2009 IEEE

an uncalibrated central catadioptric camera. It has been shown that a nonlinear optimization problem can be solved for small displacements between two images like those acquired at video rate by a camera mounted on a robot. Even though our proposed approach is similar to [1], this work is aimed at on-line self-calibration employing several views in a tracked sequence. In [1] the aim was not recovering the true intrinsic and extrinsic parameters but to align the image regions along the sequence. The proposed self-calibration algorithm needs no prior knowledge about extrinsic, camera, lens or mirror parameters. The only requirement for the user is to select a plane in the first image of a sequence. Then, the target region is warped into the reference image with respect to the sum of squared differences (SSD). In order to assess the performance of the proposed method, we perform experiments with synthetic and real data.
The structure of the paper is the following. The warping function which includes the projection model (intrinsic parameters) and the motion model (extrinsic parameters) is described in Section 2. The proposed algorithm is presented in Section 3. The results with synthetic and real data are given in Section 4. Finally, in Section 5 we made some conclusions of the work.

but to align the image regions along the sequence.

2. Warping Function

Warping is a function that allows to map the coordinates of reference image points in the current image. We will denote by $\mathbf{w}$ the warping function which depends on the homography and the sensor parameters:

$$\begin{array}{rcl} \mathbf{w}\colon \mathbb{SL}(3)\times\mathbb{R}\times\mathbb{R}^5\times\mathbb{R}^2 & \rightarrow & \mathbb{R}^2 \\ (\mathbf{H},\xi,\mathbf{K},\mathbf{p}) & \rightarrow & \mathbf{p}' = \mathbf{w}\,(\mathbf{H},\xi,\mathbf{K},\mathbf{p}) \end{array}$$

2.1. Steps of Warping and calibration parameters

The warping function include basically three transformations: *First.* The transformation between the image plane and the unit sphere. *Second*. The transformation between spheres. *Finally* The transformation between the unit sphere and the image plane.

2.1.1 The transformation between the image plane and the unit sphere

Let $\mathbf{c}^{-1}(\xi,\mathbf{K},\mathbf{p})$ be the transformation between the image plane and the unit sphere :

$$\begin{array}{rccc} \mathbf{c}^{-1}: & \mathbb{R}\times\mathbb{R}^5\times\mathbb{P}^2 & \longrightarrow & \mathbb{S}^2 \\ & (\xi,\mathbf{K},\mathbf{p}) & \longrightarrow & \mathbf{s}=\mathbf{c}^{-1}(\xi,\mathbf{K},\mathbf{p}) \end{array}$$

The transformation between the image plane and the unit sphere $\mathbf{c}^{-1}(\xi,\mathbf{K},\mathbf{p}) = \boldsymbol{\psi}^{-1}(\xi,\mathbf{k}^{-1}(\boldsymbol{\gamma},\mathbf{p}))$ lifts a point $\mathbf{p}=(u\ ,v\ ,1)^\top$ measured in the image to the unit sphere as follows :

The first step is to apply the inverse projection induced by $\mathbf{K}^{-1}$ to obtain a point on the normalized plane $\mathbf{q} = \mathbf{k}^{-1}(\boldsymbol{\gamma},\mathbf{p}) = \mathbf{K}^{-1}\mathbf{p}$. Matrix $\mathbf{K}$ is a generalised camera projection matrix (with $[f_1,f_2]^\top$ the focal length and (u_0,v_0) the principal point).

$$\mathbf{K}=\begin{pmatrix} f_1\eta & 0 & u_0 \\ 0 & f_2\eta & v_0 \\ 0 & 0 & 1 \end{pmatrix}=\begin{pmatrix} \beta_1 & 0 & u_0 \\ 0 & \beta_2 & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

A generalized camera projection matrix indicates that the sensor is no longer considered as a separate camera and mirror but as a global device. This is particularly important for calibration because it shows that f and η cannot be estimated independently. We will denote $\beta_i = f_i\eta$. In function $\mathbf{k}^{-1}(\boldsymbol{\gamma},\mathbf{p})$, $\boldsymbol{\gamma} = [\beta_1,\beta_2,u_0,v_0]^\top$ contains the camera intrinsic parameters.

The second step projects point $\mathbf{q} = (q_u,q_v,1)^\top$, on the normalized plane, using the inverse function $\alpha = \frac{-\xi-\sqrt{(1-\xi^2)(q_u^2+q_v^2)}}{q_u^2+q_v^2+1}$ proposed by Barreto [3] to obtain a point $\mathbf{s} = \boldsymbol{\psi}^{-1}(\xi,\mathbf{q}) = (\alpha q_u,\alpha q_v,\alpha+\xi)^\top$ on the unit sphere. In this step, another calibration parameter is taking into account, ξ. This parameter defines the shape of the using mirror in the omnidirectional sensor.

In this step, function $\boldsymbol{\psi}^{-1}(\xi,\mathbf{q})$ employs another calibration parameter, the mirror parameter ξ. This parameter defines the shape of the used omnidirectional mirror sensor.

Table 1. **Mirror parameter**

ξ	**Mirror type**
1	Parabolic
>0 and <1	Hyperbolic, elliptic, conical or spherical
0	Planar
>1	Fish eye
<0	No mirror

2.1.2 The transformation between spheres

Let $\boldsymbol{\hbar}(\mathbf{h},\mathbf{s})$ be the function that transforms the points between the spheres:

$$\begin{array}{rccc} \boldsymbol{\hbar}: & \mathbb{SL}(3)\times\mathbb{S}^2 & \longrightarrow & \mathbb{S}^2 \\ & (\mathbf{h},\mathbf{s}) & \longrightarrow & \mathbf{s}'=\boldsymbol{\hbar}(\mathbf{h},\mathbf{s}) \end{array}$$

The transformation between spheres $\boldsymbol{\hbar}(\mathbf{h},\mathbf{s})$ is performed by an homography.

Two points are related by an homography $\mathbf{H}$ as $\mathbf{X}' = \mathbf{HX}$. 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 $\mathbf{R}$ is the 3×3 rotation matrix of the camera and $\mathbf{t}$ its translation vector, $\mathbf{n}_d^* = \mathbf{n}^*/d^*$ 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. So, the projection of points $\mathbf{s}$ and $\mathbf{s}^{'}$, belonging to a planar region of the scene, on the sphere are related by $\rho'\mathbf{s}^{'} = \rho\mathbf{H}\mathbf{s}$. This homography has the representation of

$$\mathbf{H} = \begin{pmatrix} \mathbf{h}_1^\top \\ \mathbf{h}_2^\top \\ \mathbf{h}_3^\top \end{pmatrix} = \begin{pmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{pmatrix}$$

In this transformation, when the homography $\mathbf{H}$ is applied to point $\mathbf{s}$ a 3D point in the world is obtained by $\mathbf{X}_{3D} = \mathbf{H}\mathbf{s}$. To remind this point $\mathbf{X}_{3D}$ onto the unit sphere, a normalization is computed, $\mathbf{s}^{'} = \frac{\mathbf{X}_{3D}}{\|\mathbf{X}_{3D}\|}$. Therefore, the transformation $\boldsymbol{\hbar}(\mathbf{h}, \mathbf{s})$ is equal to : $\mathbf{s}^{'} = \boldsymbol{\hbar}(\mathbf{h}, \mathbf{s}) = \frac{\mathbf{H}\mathbf{s}}{\|\mathbf{H}\mathbf{s}\|}$. The function $\boldsymbol{\hbar}(\mathbf{h}, \mathbf{s})$ introduces the extrinsic parameters, $\mathbf{h} = [h_1, h_2, h_3, h_4, h_5, h_6, h_7, h_8, h_9]^\top$.

2.1.3 The transformation between the unit sphere and the image plane

Let $\mathbf{c}(\xi, \mathbf{K}, \mathbf{s})$ be the transformation between the sphere and the image plane:

$$\begin{array}{rccc} \mathbf{c}: & \mathbb{R} \times \mathbb{R}^5 \times \mathbb{S}^2 & \longrightarrow & \mathbb{P}^2 \\ & (\xi, \mathbf{K}, \mathbf{s}) & \longrightarrow & \mathbf{p} = \mathbf{c}(\xi, \mathbf{K}, \mathbf{s}) \end{array}$$

This transformation lifts a point $\mathbf{s}^{'} = (X_s^{'}, Y_s^{'}, Z_s^{'})^\top$ on the unit sphere to the image plane as follows : The first step is to change the point $\mathbf{s}^{'}$ to a new reference frame centered in $\mathbf{C}_p = (0, 0, \xi)$, $\mathbf{s}_{C_p}^{'} = (X_s^{'}, Y_s^{'}, Z_s^{'} + \xi)^\top$. Then the point is projected to the normalised plane to obtain $\mathbf{q}' = \boldsymbol{\psi}(\xi, \mathbf{s}^{'}) = \left(\frac{X_s^{'}}{Z_s^{'}+\xi}, \frac{Y_s^{'}}{Z_s^{'}+\xi}, 1 \right)^\top$.

The final projection involves the generalised camera projection matrix $\mathbf{K}$ to obtain the point $\mathbf{p}' = (u,\ v,\ 1)^\top$ in the image. This projection is performed as follows : $\mathbf{p}' = \mathbf{k}(\boldsymbol{\gamma}, \mathbf{q}') = \mathbf{K}\mathbf{q}'$

If expressions in sections 2.1.1, 2.1.2 and 2.1.3 are combined, the warping function can be written as:

$$\mathbf{p}' = \mathbf{w}(\mathbf{H}, \xi, \mathbf{K}, \mathbf{p}) = \mathbf{k}(\boldsymbol{\gamma}, \boldsymbol{\psi}(\xi, \boldsymbol{\hbar}(\mathbf{h}, \boldsymbol{\psi}^{-1}(\xi, \mathbf{k}^{-1}(\boldsymbol{\gamma}, \mathbf{p})))))) \quad (1)$$

This warping expression will be highly useful in the remainder of the paper.

3. Algorithm

3.1. Self-calibration Problem

The self-calibration problem will essentially be considered as an image registration problem. It is directly related to the grey-level brightness measurements in the catadioptric images. Let $\overline{\xi}$, $\overline{\mathbf{K}}$, and $\overline{\mathbf{H}}$ be the true intrinsic and extrinsic sensor parameters. The interest is to find the set of parameters $\widehat{\xi}$, $\widehat{\mathbf{K}}$, and $\widehat{\mathbf{H}}$, that fits best with the true sensors parameters such that current image will be aligned with the reference template. Therefore, when Equation 2 is satisfied, the calibration of the omnidirectional sensor is achieved.

$$I'\left(\mathbf{w}\left(\widehat{\mathbf{H}}, \widehat{\xi}, \widehat{\mathbf{K}}, \mathbf{p}\right)\right) = I(\mathbf{p}) \quad (2)$$

3.2. Method

The proposed self-calibration method is based on previous work [1]. Therefore, an image reference of the plane selected by the user is considered. This is the only requirement asked for the method to the user. No prior knowledge about extrinsic, camera, lens or mirror parameters are needed.

Let I be the selected reference image by the user. A region of size $\mathcal{R}$ (rows $\times$ columns) of I corresponding to the projection of a 3D planar region of the scene will be called *reference template*. To track the reference template in the current image $I^{'}$, we look for a set of parameters including the mirror parameter $\widehat{\xi}$, the camera intrinsic parameters $\widehat{\mathbf{K}}$ and the transformation $\widehat{\mathbf{H}}$ such that current image will be aligned with the reference template. Once we have an approximation $\widehat{\mathbf{H}}$ of the transformation $\overline{\mathbf{H}}$ and an approximation $\widehat{\xi}$ and $\widehat{\mathbf{K}}$ of the intrinsic parameters $\overline{\xi}$ and $\overline{\boldsymbol{\gamma}}$, the problem is to find the incremental transformation of $\mathbf{H}$, ξ, and $\mathbf{K}$, that minimize the sum of square differences over all the pixels of the cost function:

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

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

$$\begin{array}{rcl} \widehat{\mathbf{H}} & \leftarrow & \widehat{\mathbf{H}}\mathbf{H} \\ \widehat{\xi} & \leftarrow & \widehat{\xi} + \xi \\ \widehat{\mathbf{K}} & \leftarrow & \widehat{\mathbf{K}} + \mathbf{K} \end{array} \quad (4)$$

Similarly to [12] the incremental homography $\mathbf{H} \in \mathbb{SL}(3)$ and the intrinsic parameters updated are parametrized with local coordinates of the Lie algebra $\mathfrak{sl}(3)$.

3.3. Application of the Efficient Second-order Method (ESM)

Although the optimization problem can be solved using first-order methods, we have chosen to use the efficient second-order minimization algorithm (ESM) [11] to solve the optimization problem. In this case the least-squares solution of equation (3) is:

$$\widetilde{\mathbf{x}} = -\left(\frac{\mathbf{J}(\mathbf{0}) + \mathbf{J}(\widetilde{\mathbf{x}})}{2}\right)^{+} \mathbf{f}(\mathbf{0}) \tag{5}$$

where '+' indicates the matrix pseudo-inverse. Once the increment $\widetilde{\mathbf{x}}$ has been found, we update the estimated parameters as described in Equation (4). The current Jacobian $\mathbf{J}(\mathbf{0})$ and the approximated reference Jacobian $\mathbf{J}(\widetilde{\mathbf{x}})$ will be detailed in the next section.

3.4. Computing the current Jacobian

Let $\mathbf{x} = (\mathbf{h}, \xi, \boldsymbol{\gamma})$ be the state vector. 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. Hence, the computation of the current Jacobian $\mathbf{J}(\mathbf{0}) = \nabla \mathbf{f}_i(\mathbf{x})|_{\mathbf{x}=\mathbf{0}}$ is quite straightforward and can be decomposed into modular parts:

$$\mathbf{J}(\mathbf{0}) = \begin{bmatrix} \nabla_{\mathbf{h}} I' & \nabla_{\xi} I' & \nabla_{\boldsymbol{\gamma}} I' \end{bmatrix}_{\mathbf{x}=\mathbf{0}} \tag{6}$$

We use the chain rule to find the four derivatives of equation (6), the three blocks can be written:

$$\nabla_{\mathbf{h}} I'|_{\mathbf{x}=\mathbf{0}} = \mathbf{J}_{I'} \mathbf{J}_{\mathbf{w}} \mathbf{J}_{\mathbf{w_h}} \mathbf{J}_{\mathbf{H_h}}(\mathbf{0}) \tag{7}$$

$$\nabla_{\xi} I'|_{\mathbf{x}=\mathbf{0}} = \mathbf{J}_{I'} \mathbf{J}_{\mathbf{w}} \mathbf{J}_{\xi}(\mathbf{0}) \tag{8}$$

$$\nabla_{\boldsymbol{\gamma}} I'|_{\mathbf{x}=\mathbf{0}} = \mathbf{J}_{I'} \mathbf{J}_{\mathbf{w}} \mathbf{J}_{\boldsymbol{\gamma}}(\mathbf{0}) \tag{9}$$

$\mathbf{J}_{I'}$ represents the current image gradient.

3.5. Computing the reference Jacobian

For the reference Jacobian, it is necessary to find the current image intensity with the estimated parameters.

Similarly to the current Jacobian, the reference Jacobian $\mathbf{J}(\tilde{\mathbf{x}})$ can be computed using the derivatives chain rule. Therefore the reference Jacobian can also be written as:

$$\mathbf{J}(\tilde{\mathbf{x}}) = \begin{bmatrix} \nabla_{\mathbf{h}} I' & \nabla_{\xi} I' & \nabla_{\boldsymbol{\gamma}} I' \end{bmatrix}_{\mathbf{x}=\tilde{\mathbf{x}}} \tag{10}$$

The three blocks of this Jacobian can be written:

$$\nabla_{\mathbf{h}} I'|_{\mathbf{x}=\tilde{\mathbf{x}}} = \mathbf{J}_{I} \mathbf{J}_{\mathbf{w}} \mathbf{J}_{\mathbf{w_h}} \mathbf{J}_{\mathbf{H}^{-1}\widehat{\mathbf{H}}_{\mathbf{h}}}(\tilde{\mathbf{x}}) \tag{11}$$

$$\nabla_{\xi} I'|_{\mathbf{x}=\tilde{\mathbf{x}}} = \mathbf{J}_{I} \mathbf{J}_{\mathbf{w}} \mathbf{J}_{\xi}(\tilde{\mathbf{x}}) \tag{12}$$

$$\nabla_{\boldsymbol{\gamma}} I'|_{\mathbf{x}=\tilde{\mathbf{x}}} = \mathbf{J}_{I} \mathbf{J}_{\mathbf{w}} \mathbf{J}_{\boldsymbol{\gamma}}(\tilde{\mathbf{x}}) \tag{13}$$

In these Jacobians, $\mathbf{J}_I$ represents the gradient of the reference image, therefore, only needs to be calculated once.

3.6. Minimization procedure

The second-order incremental update is given by equation (5). However $\mathbf{J}(\tilde{\mathbf{x}})$ depends explicitly on the unknown optimal increment $\tilde{\mathbf{x}}$. By using the left invariance property $\mathbf{J}(\tilde{\mathbf{x}})\tilde{\mathbf{x}} = \mathbf{J}(0)\tilde{\mathbf{x}}$, the computation of $\mathbf{J}_{\mathbf{H}^{-1}\widehat{\mathbf{H}}_{\mathbf{h}}}(\tilde{\mathbf{x}})$, $\mathbf{J}_{\widehat{\xi}}(\tilde{\mathbf{x}})$ and $\mathbf{J}_{\widehat{\boldsymbol{\gamma}}}(\tilde{\mathbf{x}})$ are avoided by assuming $\mathbf{H} \approx \widehat{\mathbf{H}}$, $\xi \approx \widehat{\xi}$ and $\boldsymbol{\gamma} \approx \widehat{\boldsymbol{\gamma}}$. The update $\tilde{\mathbf{x}}$ of the solution can then be computed as follows:

$$\widetilde{\mathbf{x}} = \left(\left(\frac{\mathbf{J}_I + \mathbf{J}_{I'}}{2}\right) \mathbf{J}_{\mathbf{w}} \left[\mathbf{J}_{\mathbf{w_H}} \mathbf{J}_{\mathbf{H_h}}(\mathbf{0}) \;\; \mathbf{J}_{\xi}(\mathbf{0}) \;\; \mathbf{J}_{\boldsymbol{\gamma}}(\mathbf{0})\right]\right)^{+} \mathbf{f}(\mathbf{0}) \tag{14}$$

4. Results

In order to validate the proposed self-calibration method we performed experiments with synthetic data 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 $\beta_1 = -250$, $\beta_2 = -250$ and an image center $(u_0, v_0) = (512, 384)$. The homography matrices are differents for each image. The real image sequence is composed of 700 images of size 1024×768 combining a camera with a parabolic mirror. To show that the proposed method finds a good aproximation of the true intrinsic parameters, calibration of the sensor was pursued according to [13]. Determined intrinsic parameters for the real data sequence are $\xi = 1$, $\beta_1 = -295$, $\beta_2 = -294$, $u_0 = 519$ and $v_0 = 382$.

4.1. Synthetic data

Since the objective is to validate the self-calibration approach, we considered unknown intrinsic parameters to start the minimization. The initial guess for the intrinsic parameters was $\xi = 0.8$, $\beta_1 = -125$, $\beta_2 = -125$, $u_0 = 522$ and $v_0 = 394$ and for the homography parameters was given by the identity 3×3 matrix. Figure 1 shows four calibration images of the test sequence with the tracked image region marked in red. Figures 2-6 in this section show the estimated parameters during the tracking sequence. In the beginning of the estimation, the intrinsic parameters are not compute because the homography is the identity. That means that the camera has not moved, therefore, the intrinsic parameters may take any value. To avoid this, we skip a few images to ensure the camera has moved and the intrinsic parameters will take reasonable values that allow a correct convergence. As we can see the estimation is stabilized after 40 images and subsequent information does not have strong influence. The small variance of the estimates is

due to the noisy that has been added by interpolation when synthetic data has been created.

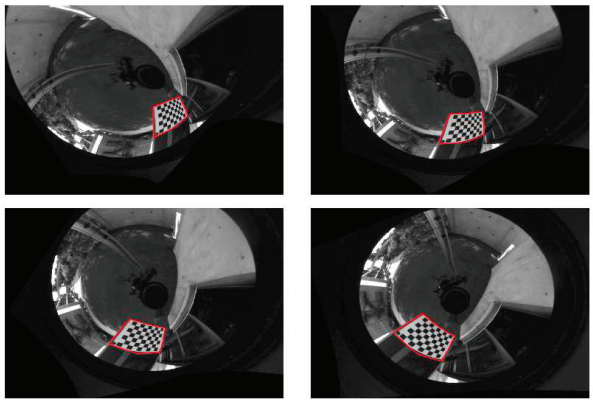

Figure 1. **Calibration**. Four calibration images of the test sequence with the tracked image region marked in red.

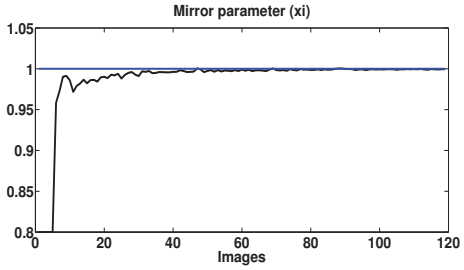

Figure 2. **Estimated parameters**. The black line depicts the estimated parameter ξ, whereas the blue line shows the ground truth parameter. The initial guess has been set to $\xi = 0.8$.

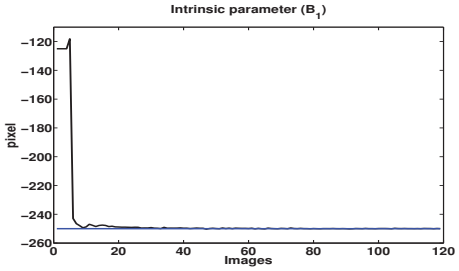

Figure 3. **Estimated parameters**. The black line depicts the estimated parameter β_1, whereas the blue line shows the ground truth parameter. The initial guess has been set to $\beta_1 = -125$.

4.2. Real data

The initial guess for the intrinsic parameters was $\xi = 0.8$, $\beta_1 = -100$, $\beta_2 = -100$, $u_0 = 512$ and $v_0 = 384$ and for the homography parameters was given by the identity

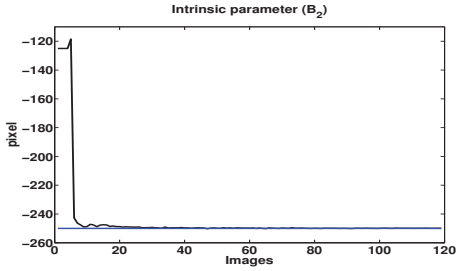

Figure 4. **Estimated parameters**. The black line depicts the estimated parameter β_2, whereas the blue line shows the ground truth parameter. The initial guess has been set to $\beta_2 = -125$.

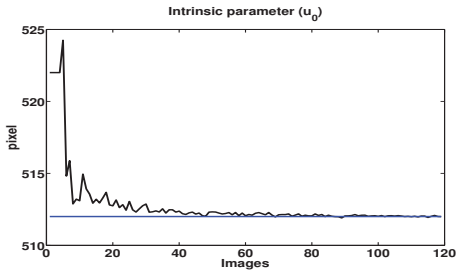

Figure 5. **Estimated parameters**. The black line depicts the estimated parameter u_0, whereas the blue line shows the ground truth parameter. The initial guess has been set to $u_0 = 522$.

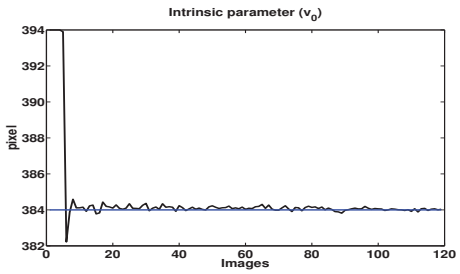

Figure 6. **Estimated parameters**. The black line depicts the estimated parameter v_0, whereas the blue line shows the ground truth parameter. The initial guess has been set to $v_0 = 394$.

3×3 matrix. Figure 7 shows four calibration images of the test sequence with the tracked image region marked in red for the proposed method in this paper. Figures 8-12 in this section show the estimated parameters during the tracking sequence. In the beginning of the estimation, the intrinsic parameters are not compute because the homography is the identity. That means that the camera has not moved, therefore, the intrinsic parameters may take any value. To avoid this, we skip a few images to ensure the camera has moved and the intrinsic parameters will take reasonable values that allow a correct convergence. To smooth the noise due to brightness and blurriness we applied Kalman filtering to the convergence curve. As we can see the estimates stabilize

after 350 images and the subsequent information does not have strong influence. The strong variance of parameters u_0 and v_0 is symptomatic of an unobservable parameter. It can be due to illumination changes and blur affect. Future work will focalise on consider illumination changes and blur affect to make more robust this method. The final calibration values are $\xi = 1.07$, $\beta_1 = 305.3$, $\beta_2 = 304.2$, $u_0 = 514.7$ and $v_0 = 377.8$.

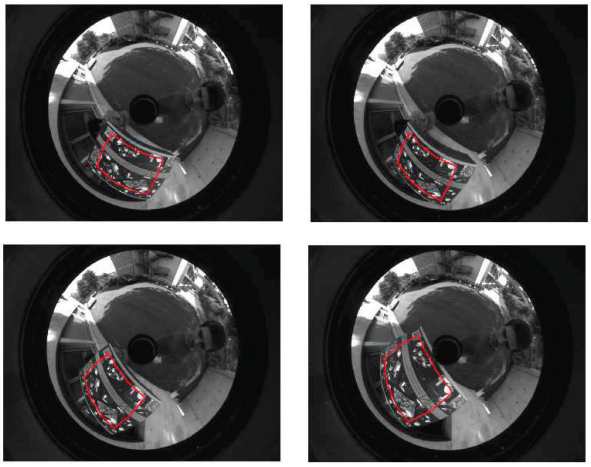

Figure 7. **Calibration**. Four calibration images of the real data with the tracked image region marked in red.

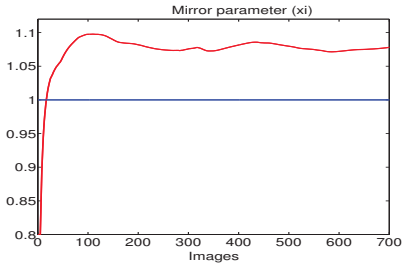

Figure 8. **Estimated parameters**. The red line depicts the estimated parameter ξ, whereas the blue line shows the ground truth parameter. The initial guess has been set to $\xi = 0.8$.

5. Conclusion

In this article, we have shown a direct approach to the self-calibration of omnidirectional cameras. The proposed self-calibration algorithm needs no prior knowledge about extrinsic, camera, lens or mirror parameters. The only requirement for the user is to select a plane in the first image of a sequence. Avoiding the awkward calibration steps should facilitate the adoption of omnidirectional sensors in robotics. Synthetic data have shown that under certainly

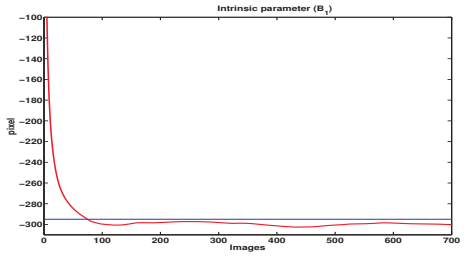

Figure 9. **Estimated parameters**. The red line depicts the estimated parameter β_1, whereas the blue line shows the ground truth parameter. The initial guess has been set to $\beta_1 = -100$.

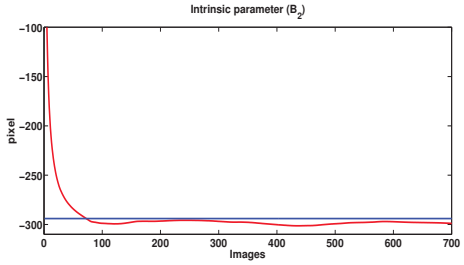

Figure 10. **Estimated parameters**. The red line depicts the estimated parameter β_2, whereas the blue line shows the ground truth parameter. The initial guess has been set to $\beta_2 = -100$.

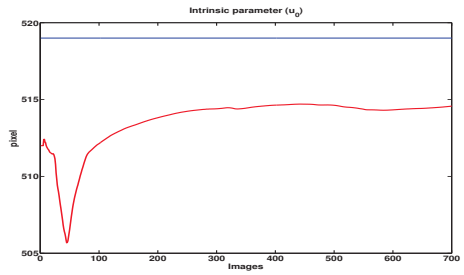

Figure 11. **Estimated parameters**. The red line depicts the estimated parameter u_0, whereas the blue line shows the ground truth parameter. The initial guess has been set to $u_0 = 512$.

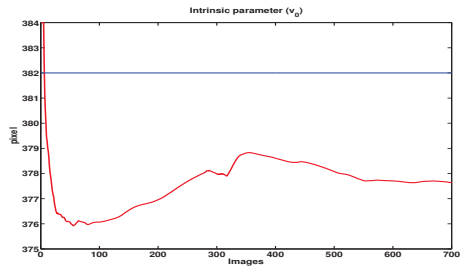

Figure 12. **Estimated parameters**. The red line depicts the estimated parameter v_0, whereas the blue line shows the ground truth parameter. The initial guess has been set to $v_0 = 384$.

conditions, no illumination changes and blur effect, the estimated parameters converges to the correct value after a few images. On other hand, experiment with real data have shown that the estimation of parameters can be unobservables because of noise. Therefore, future work will focus on considering illumination changes and blur effect to provide a faster and more robust algorithm.

References

[1] E. Malis A. Salazar-Garibay and C. Mei. Visual tracking of planes with an uncalibrated central catadioptric camera. In *IEEE International Conference on Intelligent Robots and Systems*, October 2009.

[2] Hynek Bakstein and Tomás Padjla. Panoramic mosaicing with a 180 degrees field of view lens. In *In Proceedings of the IEEE Workshop on Omndirectional Vision*, pages 60–67.

[3] J.P. Barreto and H. Araujo. Issues on the geometry of central catadioptric image formation. In *In CVPR*, 2001.

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

[5] S. Bougnoux. From projective to euclidean space under any practical situation, a criticism of self-calibration. *In Proceedings of the 6th International Conference on Computer Vision*, pages 790–796, 1998.

[6] C. Cauchois, E. Brassart, L. Delahoche, and T. Delhommelle. Reconstruction with the calibrated syclop sensor. In *IEEE International Conference on Intelligent Robots and Systems*, pages 1493–1498, Kagawa University, Takamatsu, Japan, October- November 2000.

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

[8] Christopher Geyer and Konstantinos Daniilidis. Paracatadioptric camera calibration. In *PAMI*, pages 687–695, 2002.

[9] J. Gluckman and K. Nayar. Ego-motion and omnidirectional cameras. In *In Proceedings of the IEEE ICCV'98*, pages 999–1005.

[10] S. Kang. Catadioptric self-calibration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR-00)*, pages 201–207, Los Alamitos, June 13-15 2000. IEEE.

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

[12] C. Mei, S. Benhimane, E. Malis, and P. Rives. Homography-based tracking for central catadioptric cameras. In *IEEE International Conference on Intelligent Robots and Systems*, October 2006.

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

[14] B. Micusik and T. Pajdla. Estimation of omnidirectional camera model from epipolar geometry. *In Proceedings of CVPR'03*, pages 485–490, 2003.

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