Internal Report: 2009-V
Calibration of Omnidirectional Cameras in Practice. A
Comparison of Methods\textsuperscript{1}
L. Puig, J. Bermudez and J.J. Guerrero

\textsuperscript{1}This paper was supported by projects NERO DPI2006 07928, UZ2007-TEC05 and VISPA DPI2009-14664.
Calibration of Omnidirectional Cameras in Practice.  
A Comparison of Methods

L. Puig*, J. Bermudez, J.J. Guerrero 
I3A-DIIS, Maria de Luna 1, Zaragoza, Spain 

Abstract

Omnidirectional cameras are becoming increasingly popular in computer vision and robotics. Camera calibration is a step before performing any task involving metric scene measurement, required in nearly all robotics tasks. In recent years many different methods to calibrate central omnidirectional cameras have been developed, based on different internal models and many times some of them limited for a specific mirror shape. In this paper we review the existing methods designed to calibrate any central omnivision system and analyze their advantages and drawbacks doing a deep practical comparison. We choose methods available as OpenSource and which do not require a difficult pattern or scene. The evaluation method of calibration accuracy also considers 3D metric reconstruction combining omnidirectional images. Comparative results are shown and discussed in detail.

1. Introduction

In recent years the use of omnidirectional cameras among the Robotics community has widely increased. One of the major advantages of this kind of cameras is its wide field of view (FOV) which allows them to have a view of the whole scene. There exist several types of omnidirectional cameras. We can find the rotating camera, which is a conventional camera with a mechanic system that allows it to move along a circular trajectory and to acquire images from the surroundings. Arrays of cameras are sets of conventional cameras pointing to different directions in a circular configuration. The dioptric systems, which are conventional cameras with wide-angles lenses, like the fish-eye lenses. One more type to which we give special attention is the catadioptric systems. This type of systems combine conventional cameras with mirrors. The catadioptric

*This paper was supported by projects NERO DPI2006 07928, UZ2007-TECO5 and VISPA DPI2009-14664.
*Corresponding author

Email addresses: lpuig@unizar.es (L. Puig), jbermudez@unizar.es (J. Bermudez), jguerrer@unizar.es (J.J. Guerrero)
systems were studied by Baker and Nayar [1]. They proved that the elliptic, parabolic and hyperbolic mirrors are the only ones that can be combined with conventional cameras and provide the single-viewpoint property. The most popular of such systems are the hyper-catadioptric one which is composed by a hyperbolic mirror and a perspective camera and the para-catadioptric system which is composed by a parabolic mirror and an orthographic camera. The use of these systems can be seen in applications such as surveillance, localization, virtual reality, navigation, SLAM, visual odometry, etc. Many of these tasks, specially in robotics, require to recover metric information about the environment since the robot has to interact in a real scenario where 3D information is crucial. This metric information depends on the complete calibration of the catadioptric system.

There exist several geometric and analytic models to deal with catadioptric omnidirectional systems. Svoboda and Pajdla [2] propose different models for different mirrors. Strelow et al. [3] deal directly with the reflection properties of the rays on the mirror. A unified model was proposed by Geyer and Daniilidis [4]. In this work they present the sphere camera model which allows to deal with any central catadioptric system. Later this model was extended by Barreto and Araujo [5]. This sphere model is the most used model in current days. Another category that should be mentioned is the generic methods that can calibrate any arbitrary imaging system. Grossberg and Nayar [6] propose a method based on virtual sensing elements called raxels which describe a mapping from incoming scene rays to photo-sensitive elements on the image detector. This work has inspired many works and a list of some of them can be found in Ramalingam et al. [7].

Recently many calibration methods for central catadioptric systems have been developed. Most of them deal with the catadioptric system as a whole. Some others separate the calibration of the perspective/orthographic camera from the computing of the mirror parameters [2, 8]. Some of these methods are also valid for fish-eye lenses. Among the calibration methods which consider the whole catadioptric system we can mention the ones which use single or multiple views of a 2D pattern [9, 10, 11, 12, 13], 3D pattern [14], cylinder pattern [15], the ones which use a single image containing invariant features like lines [16, 17, 18, 19, 20, 21, 22, 23, 24, 25], self-calibration approaches [26, 27].

As we can see the amount of calibration methods is considerable. Some works have tried to either classify or to compare them. In [20] a rough classification of the methods in three categories is performed. A more detailed classification is given in [12] where five categories are proposed. These classifications are out of date and miss some methods. Besides that they do not perform a comparison of available approaches. Frank et al. [13] perform a rough comparison of their method to [10] and [11]. However, they simply use the reprojection error criteria. Toepfer and Ehlen[15] perform another comparison between their method and [10, 11] but the comparison criteria used does not give a clear idea of the behavior of these approaches. So up to our knowledge there no exist a classification and comparison of calibration methods like the one proposed in this paper.

In robotics applications the calibration of the sensors is mandatory to per-
form many tasks. It allows the robot to acquire metric information from the environment. Several calibration methods are proposed but just a few are available as OpenSource and with easy usage requirements. These reasons are important from a practical point of view like required by a robotician user, usually interested in 3D motion and structure results and non interested in complex projection models. To compare these methods we perform a 3D reconstruction where metric information is computed. These compared methods are:

1. Mei and Rives [11] which use the sphere camera model and requires several images of a 2D pattern. We will call this approach Sphere-2DPattern.
2. Bastanlar et al. [14] which obtains a solution in closed form requiring a set 3D-2D correspondences. It also uses the sphere camera model. We call this approach DLT-like.
3. Barreto and Araujo [18] uses also the sphere camera model and requires a single omnidirectional image containing a minimum of three lines. We call it Sphere-Lines.
4. Scaramuzza et al. [10] which models the omnidirectional images as distorted images where the parameters of distortion have to be found. We call this approach Distortion-2DPattern.

This paper is divided as follows. Section 2 presents a review of the existing calibration methods up to now. Section 3 explains the sphere camera model, which is used by three of the methods in this comparison. Section 4 briefly presents the four methods enunciated above. In Section 5 experiments are performed and commented. Finally Section 6 enunciates the conclusions.

2. Existing Calibration Methods for Omnidirectional Systems

As observed above there exist many calibration methods. They use different techniques and models to calibrate the omnidirectional systems. Some of them first calibrate the perspective camera and finding after that the mirror parameters.

   Line-based calibration. Many methods are based on the projection of lines in the omnidirectional images. The main advantage of using lines is that they are present in many environments and a special pattern is not needed. These approaches compute the image of the absolute conic from which they compute the intrinsic parameters of the catadioptric system. Geyer and Daniilidis [16] calibrate para-catadioptric cameras from the images of only three lines. Barreto and Araujo [19] study the geometric properties of line images under the central catadioptric model. They give a calibration method suitable for any kind of central catadioptric system. Ying and Hu [20] analyze the relation of the camera intrinsic parameters and sphere imaged contour. They use the projection of lines as well as projections of the sphere. The former gives three invariants and the latter two. Ying and Zha [21] show that all line images from a catadioptric camera must belong to a family of conics which is called a line image family related to certain intrinsic parameters. They present a Hough transform
for line images detection which ensures that all detected conics must belong to a line image family related to certain intrinsic parameters. Vanderportaele et al. [22] slightly improves [16] using a geometric distance instead of an algebraic one and they allow to deal with lines that are projected to straight lines or to circular arcs in an unified manner. Wu et al. [23] introduce a shift from the central catadioptric model to the pinhole model from which they establish linear constraints on the intrinsic parameters. Without conic fitting they are able to calibrate para-catadioptric-like cameras. Caglioti et al. [24] use the image of one generic space line, from which they derive some constraints that, combined with the harmonic homology relating the apparent contours of the mirror allow them to calibrate the catadioptric system. More recently Wu et al. [25] derive the relation between the projection on the viewing sphere of a space point and its catadioptric image. From this relation they establish linear constraints that are used to calibrate any central catadioptric camera. Vasseur and Mouaddib [28] detect lines in the 3D scene which are later used to calculate the intrinsic parameters. This approach is valid for any central catadioptric system.

**2D pattern calibration.** This kind of methods use a 2D calibration pattern with control points. These control points can be corners, dots, or any features that can be easily extracted from the images. Using iterative methods extrinsic and intrinsic parameters can be recovered. Scaramuzza et al. [10] propose a technique to calibrate single viewpoint omnidirectional cameras. They assume that the image projection function can be described by a Taylor series expansion whose coefficients are estimated by solving a two-step least squares linear minimization problem. Mei and Rives [11] propose as Scaramuzza a flexible approach to calibrate single viewpoint sensors from planar grids, but based on an exact theoretical projection function -the sphere model- to which some parameters as distortion are added to consider real-world errors. Deng et al. [12] use the bounding ellipse of the catadioptric image and the field of view (FOV) to obtain the intrinsic parameters. Then, they use the relation between the central catadioptric and the pinhole model to compute the extrinsic parameters.

**3D Point-based calibration.** These methods require the position of 3D points observed usually in a single image. Aliaga [29] propose an approach to estimate camera intrinsic and extrinsic parameters, where the mirror center must be manually determined. This approach only works for para-catadioptric systems. Wu and Hu [30] introduced the invariants of 1D/2D/3D space points and use them to compute the camera principal point with a quasi-linear method. An approach based on the Direct Linear Transformation (DLT) using lifted coordinates to calibrate any central catadioptric camera is proposed in [14]. It computes a generic projection matrix valid for any central catadioptric system. From this matrix the intrinsic and extrinsic parameters are extracted in a closed form and refined by non-linear optimization afterwards. This approach requires a single omnidirectional image containing points spread in at least three different planes.

**Self-calibration.** This kind of calibration techniques uses only point correspondences in multiple views, without needing to know either the 3D location of the points or the camera locations. Kang [26] uses the consistency of pair-
wise tracked point features for calibration. The method is only suitable for para-catadioptric systems. Micusik and Pajdla [27] propose a method valid for fish-eye lenses and catadioptric systems. They show that epipolar geometry of these systems can be estimated from a small number of correspondences. They propose to use a robust estimation approach to estimate the image projection model, the epipolar geometry and to avoid outliers.

**Polarization Imaging**. This method is proposed by Morel and Fofi [8]. It is based on an accurate reconstruction of the mirror by means of polarization imaging. It uses a very simple camera model which allows them to deal with any type of camera. However they observe that developing an efficient and easy-to-use calibration method is not trivial.

In Table 1 we present and classify the existing approaches to calibrate omnidirectional catadioptric systems, according to different criteria: required pattern, number of views, analytical model, type of mirror, or if the method requires the separated calibration of the mirror and the camera.

### 3. Sphere Camera model

As many of the calibration methods are based on the sphere camera model [4, 5], we explain this model here, and later we show the modifications added in each approach.

All central catadioptric cameras can be modeled by a unit sphere and a perspective camera, such that the projection of 3D points can be performed in two steps (Fig. 1). First, one projects the point onto the sphere, to the intersection of the sphere and the line joining its center and the 3D point. There are two such intersection points, \( r_± \). These points are then projected into a normalized plane \( π_{nor} \) resulting in two points, \( q_± \). Finally these points are projected into the perspective plane \( π_p \) giving again two points \( x_± \), one of which is physically true. This model covers all central catadioptric cameras, encoded by \( ξ \), which is the distance between the perspective camera and the center of
the sphere. \( \xi = 0 \) for perspective, \( \xi = 1 \) for para-catadioptric, \( 0 < \xi < 1 \) for hyper-catadioptric. This parameter can be computed based on the mirror information. Let the unit sphere be located at the origin and the optical center of the perspective camera, at the point \( C_p = (0, 0, -\xi)^T \). The perspective camera is modeled by the projection matrix \( P \sim A_p R_p (I - C_p) \), where \( A_p \) is its calibration matrix. The rotation \( R_p \) denotes a rotation of the perspective camera looking at the mirror. Since both intrinsic and extrinsic parameters of the perspective camera are intrinsic parameters for the catadioptric camera, we replace \( A_p R_p \) by a generic projective transformation \( H_c \). Note that the focal length of the sphere model is a value determined by the actual camera focal length and the mirror shape parameters \( (\xi, \psi) \) which are computed from the mirror parameters \([19, 11]\). The intrinsic parameters of the catadioptric camera are thus \( \xi \) and \( H_c \), where \( H_c \) is defined as

\[
H_c = A_p R_p \begin{bmatrix}
\psi - \xi & 0 & 0 \\
0 & \xi - \psi & 0 \\
0 & 0 & 1
\end{bmatrix}
\]  

Observe the change of the sign in the elements of the diagonal of \( M_c \). This encodes the mirror effect which causes a flip on the omnidirectional image. Mei et al. \([11]\) just take into account the difference \( \xi - \psi \) and call it \( \eta \), so the flip is not considered in that model. Following that, Eq. 2 encodes this two projections (\( x_\pm \)) and its corresponding inverse function is Eq. 3, which maps an image point \( x \) to an oriented 3D ray \( X \).

\[
x_\pm = h(X) = \begin{bmatrix} X \\ Y \\ Z \pm \xi \sqrt{X^2 + Y^2 + Z^2} \end{bmatrix}
\]  

\[
h^{-1}(x) = \begin{bmatrix} \xi + \sqrt{1 + (\frac{(-\xi^2)(x^2 + y^2)}{x^2 + y^2 + 1})} x \\ \xi + \sqrt{1 + (\frac{(-\xi^2)(x^2 + y^2)}{x^2 + y^2 + 1})} y \\ \xi + \sqrt{1 + (\frac{(-\xi^2)(x^2 + y^2)}{x^2 + y^2 + 1})} - \xi \\ \end{bmatrix}
\]

Some authors \([19, 11]\) just take into account the positive projection given by function \( h \) but some others \([31, 32]\) use lifted coordinates to deal with the non-linearities present in function \( h \). The vector lifting consists in mapping a 3-vector in the projective space \( \mathbb{P}^2 \) to a 6-vector in the projective space \( \mathbb{P}^5 \) and its performed by the following equation

\[
\hat{x} = (x_1^2 \ x_1 x_2 \ x_2^2 \ x_1 x_3 \ x_2 x_3 \ x_3^2)^T
\]  

Lifted matrices are also used to compute the generic catadioptric projection matrix. Two different ways to compute this lifting can be found in \([14, 31]\).
4. Calibration Methods analyzed

In this section we summarize the four OpenSource methods used to calibrate the omnidirectional systems. The purpose of this section is just to show a general view of the methods. If the reader wants to know more about a particular method, the corresponding reference should be checked.

4.1. Sphere-2DPattern approach

This approach [11] uses the sphere model explained in the last section, with the difference that it does not consider the image flip induced by \((\psi - \xi)\), it uses \((\xi - \psi)\) in \(x\) and \(y\) coordinates. This approach adds to this model, distortion parameters to consider real world errors. This method is multiview, which means that it requires several images of the same pattern containing as much points as possible. This method needs the user to provide prior information to initialize the principal point and the focal length of the catadioptric system. The principal point is computed from the mirror center and the mirror inner border. The focal length is computed from three or more collinear non-radial points. Once all the intrinsic and extrinsic parameters are initialized a non-linear process is performed. This approach is also valid for fish-eye lenses and spherical mirrors.

4.1.1. Projection model

The author enlists 17 parameters involved in the projection of a 3D point \(X\) into one image point \(x = (x, y)\). These parameters are organized in vectors \(V^i\) and noted by transformations as follows

1. \((P_1)\). Extrinsic parameters. \(V^1 = [q_1, q_2, q_3, q_4, t_1, t_2, t_3]\) where quaternions are used to compute the rotation between the camera reference system and the world reference system.
2. (P_2). Mirror transformation. $V^2 = [\xi]$. This consists in applying Eq. 2 that depends only on $\xi$.

3. (P_3). Distortion. $V^3 = [k_1, k_2, p_1, p_2]$. The sources of distortion are the well known radial and tangential distortion [33]. These distortions are applied before the collineation induced by $[\psi - \xi]$. Two coefficients are used for each kind of distortion.

4. (P_4). Camera model. $V^4 = [\theta \gamma_1 \gamma_2 u_0 v_0]$. A standard pin-hole model is used for the generalized camera projection $A_p$, where $(\gamma_1, \gamma_2)$ are the focal lengths of the catadioptric system for $x$ and $y$ axis, $\theta$ is the skew parameter, and $(u_0, v_0)$ is the principal point.

### 4.1.2. Minimization Criteria

Combining all the projection functions into $G$ and $V$ being the 17 parameters of the intrinsic and extrinsic calibration:

$$G = P_1 \circ P_2 \circ P_3 \circ P_4, \quad V = [V^1 V^2 V^3 V^4]$$

If the pattern used to calibrate the camera is composed of $m$ points $X_i$, with their associated image values $x_i$, the solution of the calibration problem can be obtained by minimizing the Euclidean distance between the projection of the pattern and the extracted values in the image. This cost function is the following:

$$F(x) = \frac{1}{2} \sum_{i=1}^{m} [G(V, X_i) - x_i]^2$$

The non-linear approach used to minimize Eq. 6 is the Levenberg-Marquardt.

### 4.2. Sphere-Lines approach

This method [19] is based on the computing of the absolute conic and the mirror parameter under the sphere camera model. In omnidirectional images 3D lines are mapped to conics. So the first step is to fit these conics. With the information provided by these conics and the location of the principal point an intermediate step is performed. It computes entities like polar lines, lines at infinity and circle points. From these intermediate entities and some invariant properties like collinearity, incidence and cross-ratio the image of the absolute conic and the mirror parameter are computed.

#### 4.2.1. Calibration of the Central Catadioptric Camera

To calibrate the central catadioptric camera we have to compute the image of the absolute conic $\Omega_{\infty} = H_c^{-1} H_c^{-1}$ and the mirror parameter $\xi$. From the image of a conic in the omnidirectional image is possible to compute two points that lie on the image of the absolute conic. Since a conic is defined by a minimum of 5 points at least three conic images are required to obtain $\Omega_{\infty}$. Once the image of the absolute conic is computed $H_c$ is determined from the Cholesky decomposition of $\Omega_{\infty}$.

The mirror parameter $\xi$ is computed following these steps:
1. For each pair of conics \( \hat{\Omega}_i, \hat{\Omega}_j \) determine the corresponding line which passes through the poles of \( \hat{\mu}_{ij} \) with respect to \( \hat{\Omega}_i, \hat{\Omega}_j \). \( \hat{\mu}_{ij} \) lies on line \( \hat{P}_{i\infty} \).

2. Each line image \( \hat{\Omega}_i \) has a pencil of lines associated with it. Determine the point \( \hat{N}_i \) where those lines intersect.

3. Determine the line \( \hat{\mu}_i = \hat{N}_i \land \hat{O} \) associated with the catadioptric line image \( \hat{\Omega}_i \).

4. Estimate the locus of the line at infinity \( \hat{\pi}_\infty \) knowing that it goes through the poles of \( \hat{\mu}_i \) with respect to conic \( \hat{\Omega}_i \).

5. For each conic \( \hat{\Omega}_i \) determine the intersection point \( \hat{D}_i = \hat{\mu}_i \land \hat{\pi}_i \) and the pole \( \hat{C}_i = \hat{\Omega}_i \land \hat{\pi}_\infty \) which must lie on line \( \hat{\mu}_i \).

6. The mirror parameter \( \xi \) is provided by the square root of the cross ratio between the points \( \hat{O}, \hat{D}_i, \hat{N}_i \) and \( \hat{C}_i \).

For a more detailed description of this approach [19] should be consulted.

### 4.3. DLT-like approach

This approach [14] also uses the sphere camera model. To deal with the non-linearities present in this model, the lifting of vectors and matrices is used. This method computes a lifted \( 6 \times 10 \) projection matrix that is valid for all single-viewpoint catadioptric cameras. The required input for this method is a single image with a minimum of 20 3D-2D correspondences distributed in 3 different planes.

#### 4.3.1. Generic Projection Matrix

A 3D point \( \mathbf{X} \) is mathematically projected to two image points \( \mathbf{x}_+, \mathbf{x}_- \), which are represented via the degenerate dual conic \( \Omega \). The relation between them is

\[
\Omega \sim \mathbf{x}_+ \mathbf{x}_+^\top + \mathbf{x}_- \mathbf{x}_-^\top
\]

This conic represented as a 6-vector \( \mathbf{c} = (c_1, c_2, c_3, c_4, c_5, c_6)^\top \) projected on the omnidirectional image is computed using the lifted 3D point coordinates, intrinsic and extrinsic parameters

\[
\mathbf{c} \sim \begin{pmatrix} \hat{H}_{[6 \times 6]} & \hat{R}_{[6 \times 6]} \end{pmatrix} \mathbf{X}_\xi \begin{pmatrix} \mathbf{I}_{6 \times 4} \\
T_{6 \times 4} \end{pmatrix} \hat{X}_{10}
\]

where, \( \hat{R} \) represents the rotation of the catadioptric camera. \( X_\xi \) and \( T_{6 \times 4} \) depend only on the sphere model parameter \( \xi \) and position of the catadioptric camera \( \mathbf{C} = (t_x, t_y, t_z) \) respectively. Thus, a \( 6 \times 10 \) catadioptric projection matrix, \( P_{cata} \), is expressed by its intrinsic \( A_{cata} \) and extrinsic \( T_{cata} \) parameters

\[
P_{cata} = \begin{pmatrix} \begin{pmatrix} \hat{H}_{[6 \times 6]} & \hat{R}_{[6 \times 6]} \end{pmatrix} \\
A_{cata} & T_{cata} \end{pmatrix}
\]

This matrix is computed from a minimum of 20 3D-2D lifted correspondences in a similar way than the perspective case [34] using least squares.
The 60-vector $p_{\text{cata}}$ contains the 60 coefficients of $P_{\text{cata}}$. Manipulating this matrix algebraically the intrinsic and extrinsic parameters are extracted. These extracted values are used as an initialization to perform a non-linear process (Levenberg-Marquardt). In this process some parameters that are not included in the sphere camera model are added. These parameters are, as in Sphere-2DPattern approach, the radial and tangential distortion which are initialized to zero. This approach use two parameters for each type of distortion. The minimization criterion is the root mean square (RMS) of distance error between a measured image point and its reprojected correspondence.

4.4. Distortion-2DPattern approach

In this approach [10] the only assumption is that the image projection function can be described by a Taylor series expansion whose coefficients are estimated by solving a two-step least squares linear minimization problem. It does not require either any a priori knowledge of the motion or a specific model of the omnidirectional sensor. So, this approach assumes that the omnidirectional image is in general a high distorted image and we have to compute the distortion parameters to obtain such a distorted image. This approach as Sphere-2DPattern requires several images from different position of a 2D pattern. The accuracy depends on the number of images used and on the degree of the polynomial.

4.4.1. Omnidirectional Camera Model

Let $X$ be a scene point and $x'' = \begin{bmatrix} x''_1 \\ y''_1 \end{bmatrix}$ be its projection onto the sensor plane and $x' = \begin{bmatrix} x'_1 \\ y'_1 \end{bmatrix}$ its image in the camera plane. The two systems are related by an affine transformation, which is substituted by a scale factor $\alpha$ giving the relation $x''_1 = \alpha x'_1$. The projection function $g$, captures the relation between a point $x''$ in the sensor plane and the vector $p$ emanating from the viewpoint $O$ to the scene point $X$. So, the complete model of an omnidirectional camera is

$$\lambda p = \lambda g(x'') = \lambda g(\alpha x') = \lambda g(\alpha x') = P_X, \quad \lambda > 0,$$

(10)

where $X \in \mathbb{R}^4$ is expressed in homogeneous coordinates; $P \in \mathbb{R}^{3 \times 4}$ is the perspective projection matrix. The function $g$ has the following form

$$g(x') = (x', y', f(x', y'))^T,$$

(11)

and $f$ is defined as

$$f(x', y') = a_0 + a_1 \rho' + a_2 \rho'^2 + \cdots + a_n \rho'^n$$

(12)

where $\rho'$ is the Euclidean distance between the image center and the point. As $g$ is a linear function the scalar $\alpha$ is taken out. Then $g(\alpha u') = \alpha g(u')$. The factor
\( \alpha \) is directly integrated in the depth factor \( \lambda \). Thus, in order to calibrate the omnidirectional camera the \( n + 1 \) parameters \((a_0, a_1, a_2, \ldots, a_n)\) corresponding to the coefficients of function \( f \) need to be estimated.

**4.4.2. Camera Calibration**

During the calibration procedure, a planar pattern of known geometry is shown at different unknown positions. The camera calibration under this approach is performed in two steps. The first step is to compute the extrinsic parameters, i.e., the relation between each location of the planar pattern and the sensor coordinate system. Each point on the pattern gives three homogeneous equations. Only one of them is linear and it is used to compute the extrinsic parameters. In the second step, the intrinsic parameters are estimated, using the extrinsic parameters previously computed and the other two equations. After these two steps the camera has been calibrated. The authors do not mention it, but after this linear process a non-linear optimization is performed using the Gauss-Newton algorithm\(^2\).

**5. Experiments**

In order to compare the different calibration methods we calibrate a hyper-catadioptric system\(^3\), a fish-eye and a commercial proprietary shape mirror\(^4\). We call this system unknown-shape. As an additional non-central system we move the perspective camera of the hyper-catadioptric system far from the mirror. This cause that the optical center of the perspective camera is not located in the other focus described by the hyperbolic mirror. We calibrate the omnidirectional systems with the four methods and compare the results with a reconstruction experiment which is explained in section 5.1. The set up used to calibrate the omnidirectional system for every method is explained as follows.

*Sphere-2DPattern approach.* This approach requires images of a single 2D pattern. These images have to cover most of the omnidirectional image area. This approach asks the user for the image center and for the mirror inner border to compute the principal point. Then asks for four aligned edge points on a non-radial line to compute the focal length. With this information it asks for the four point corner and uses a subpixel technique to extract the rest of the points present in the pattern. If the focal length estimation is not well estimated then all points have to be given manually.

*DLT-like approach.* In the DLT-like approach a single image of a 3D pattern was used. This approach does not have an automatic extractor so all points are given manually. This method requires as input a set of points lying on three different planes.

\(^2\)This algorithm is provided by the lsqnonlin Matlab function

\(^3\)Neovision H3S with XCD-X710 SONY camera

\(^4\)http://www.0-360.com
Sphere-Lines approach. This approach is based on the projections of lines in the omnidirectional images. This method only requires one omnidirectional image containing at least three lines.

Distortion-2D approach. This approach also uses images coming from a 2D pattern. The last version of this method has an automatic corner detector which detects most of the corners present in a single pattern. The amount of corners given manually is minimum. Once all the points in all the images are given the calibration process starts.

5.1. Evaluation Criteria

In order to compare the different approaches used to calibrate central catadioptric cameras we perform a task where a calibrated camera is required. We decide to avoid the reprojection error given by each method when the calibration is performed because this reprojection error is obtained from the same images used for the calibration. Instead we choose the Structure from Motion from two calibrated omnidirectional images of a 3D pattern (Fig. 2(a)) built in our lab. The pattern has been measured with accuracy using photogrammetric software\(^5\). Thus, a 3D reconstruction by bundle adjustment has been made. We use 6 convergent views taken with a calibrated high-resolution camera (Canon EOS 5D with 12.8Mpix.). The estimated accuracy of the location of the 3D points is better than 0.1mm. Fig. 2(b) shows the configuration used for the SfM experiment. Using the calibration provided by each method we compute the corresponding 3D rays from each omnidirectional image. We used these correspondences of 3D rays to compute the essential matrix $E$ which relates them. From this matrix we compute two projection matrices $P_1 = [I|0]$ and $P_2 = [R|t]$. Then, with these projection matrices and the 3D rays we compute an initial

---

\(^5\)Photomodeler software was used.
3D reconstruction using a linear triangulation method [34] which is later refined by a bundle adjustment optimization process. The 3D reconstruction depends on the number of correspondences. We use a set of 144 points to compute the reprojection error and to evaluate the 3D reconstruction results. All the correspondences were given manually. We choose two different criteria to measure the accuracy of each model. These criteria are:

- The average error between the real 3D points and their estimations.
- The reprojection error. We project the ground truth 3D pattern in the two cameras with the locations given by the SFM algorithm. We measure the error in pixels between the image points and the ground truth reprojection.

5.2. Hyper-catadioptric System

The hyper-catadioptric system is composed by a perspective camera with a resolution of 1024 × 768 and a hyperbolic mirror having a 60 mm diameter and parameters $a = 281$ mm and $b = 234$ mm according to manufacturer information. The mirror parameter for the sphere camera model is $\xi = 0.9662$. In Fig. 3 we observe some of the images used to calibrate this system. We use 11 images of the 2D pattern to calibrate the system with the Sphere-2DPattern and Distortion-2D approaches.

5.2.1. Mirror parameter and principal point

Three of these methods are based on the sphere camera model. In Table 2 we present the estimations of the principal point $(u_0, v_0)$ and the mirror parameter
Table 2: Comparison of the physical parameters given by the 3 methods based on the sphere model.

<table>
<thead>
<tr>
<th>Method</th>
<th>( \xi )</th>
<th>((u_0, v_0))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manufacturer Info</td>
<td>0.9662</td>
<td>(511.88, 399.25)</td>
</tr>
<tr>
<td>Sphere-2DPattern</td>
<td>0.9684</td>
<td>(513.93, 400.76)</td>
</tr>
<tr>
<td>Sphere-Lines</td>
<td>1.0215</td>
<td>(523.82, 416.29)</td>
</tr>
<tr>
<td>DLT-like</td>
<td>0.9868</td>
<td>(509.95, 398.54)</td>
</tr>
</tbody>
</table>

Table 3: Accuracy of 3D coordinate measurement in mm and reprojection error in pixels for the hyper-catadioptric system.

<table>
<thead>
<tr>
<th>Method</th>
<th>3D Error (mm)</th>
<th>Reprojection Error (pixels)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>( \sigma )</td>
</tr>
<tr>
<td>Sphere-2DPattern</td>
<td>1.41</td>
<td>0.83</td>
</tr>
<tr>
<td>DLT-like</td>
<td>1.03</td>
<td>0.56</td>
</tr>
<tr>
<td>Sphere-Lines</td>
<td>2.01</td>
<td>0.98</td>
</tr>
<tr>
<td>Distortion-2DPattern</td>
<td>1.27</td>
<td>0.78</td>
</tr>
</tbody>
</table>

\( \xi \) since they are related with sphere model parameters. Distortion-2DPattern model does not offer information about the catadioptric system. As we can see, the best estimation of the mirror parameter is given by Sphere-2DPattern but also the DLT-like algorithm gives a good approximation. Sphere-Lines gives a value bigger than 1 which does not correspond to a hyperbolic mirror. With respect to the principal point the estimation of Sphere-2DPattern and DLT-like are close to the real one. The difference is that Sphere-2DPattern method asks the user to give the image center and the rim of the mirror to compute the principal point and DLT-like algorithm does not need any of this a priori information.

In Table 3 we observe the accuracy in millimeters of the reconstruction of the 3D pattern and the reprojection error given by each of the approaches. The best reconstruction was obtained with the calibration given by the DLT-like algorithm with an average error of 1.03 \( \text{mm} \). We also observe the reprojection error in pixels where the smallest error corresponds to DLT-like and Distortion-2DPattern with 0.3 pixels.
Figure 5: Some images used to calibrate the fish-eye system. (a) Sphere-Lines. (b) DLT-like approach. (c, d) Sphere-2DPattern and Distortion-2DPattern approaches.

5.3. Fish-eye Lens

The fish-eye lens used in this experiment is a Raynox DCR-CF185PRO with a FOV of 185\(^\circ\) on all directions. It is mounted on a high-resolution camera. In Fig. 3 we show some of the images used to calibrate this system. We use 8 images to perform the calibration with the Sphere-2DPattern and Distortion-2DPattern approaches. We use the image of 7 lines to perform the calibration with the Sphere-Lines approach. As none of the methods give any information about the system we just show the results obtained from the SfM experiment.

In Table 4 we show the results of this experiment. We observe that the methods that claim that are able to calibrate or at least estimate the non-central fish eye systems give good results. This is the case of the Sphere-2DPattern, DLT-like and the Distortion-2DPattern approaches. The opposite case is the Sphere-Lines approach which is not able to correctly calibrate the system. It gives reprojection errors close to the one hundred pixels. We observe that the behavior of the approaches that can calibrate the system give similar results. The DLT-like gives the best results with a reprojection error smaller than 2 pixels and an average 3D error smaller than 1 mm. We observe that this approach gives the highest maximum individual error (22.15 mm). This is explained because the lowest point in one of the images (see Fig. 6(a)) is very close to the border and the calibration image has no points close to the border.
Table 4: Accuracy of 3D coordinate measurement in mm and reprojection error in pixels for the fish-eye lens.

<table>
<thead>
<tr>
<th></th>
<th>3D Error (mm)</th>
<th>Reprojection Error (pixels)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>σ</td>
</tr>
<tr>
<td>Sphere-2DPattern</td>
<td>1.12</td>
<td>0.90</td>
</tr>
<tr>
<td>DLT-like</td>
<td>0.82</td>
<td>1.86</td>
</tr>
<tr>
<td>Distortion-2DPattern</td>
<td>1.49</td>
<td>1.19</td>
</tr>
</tbody>
</table>

5.4. Unknown-shape Catadioptric System

This system is the combination of a commercial proprietary shape mirror and a high-resolution camera. As a catadioptric system any of the methods analyzed here can calibrate it. We use 6 images of a 2D pattern to calibrate this system for the Sphere-2DPattern and Distortion-2DPattern approaches. We use an image with 4 lines to perform the calibration with the Sphere-Lines approach. We observed several difficulties when more than 4 lines were used to calibrate the system using the Sphere-Lines approach. Sometimes the calibration was given using complex numbers or it had problems of convergence. In Fig. 7 we observe some of the images used to calibrate this system under all the analyzed approaches. In Fig. 8 we show the images used to perform the SfM experiment.

We decided to use a more complex reconstruction scenario to observe the behavior of the approaches under these conditions (Fig. 8 (a), (b)). The results of the experiments where the 3D accuracy and the reprojection error are shown in Table 5. We observe that the smallest error is given by the Distortion-2DPattern with a 3D error of 2.61 mm and a reprojection error of 0.32 pixels. Again we observe a similar behavior where a distance of two and four millimeters separate the second and the third method, respectively. The reprojection error for all the methods is below 2 pixels.

5.5. Non-central hyper-catadioptric system

This system is the same described in the hyper-catadioptric experiment. The only difference is that the perspective camera is displaced as far as possible.
Figure 7: Some images used to calibrate the unknown-shape catadioptric system. (a) Sphere-Lines. (b) DLT-like approach. (c,d) Sphere-2DPattern and Distortion-2DPattern approaches.

Figure 8: Images used in the SfM experiment with reprojected points superimposed (unknown-shape).

from the mirror. This cause that the optical center of the perspective camera is not located at the other focus described by the hyperbolic mirror. The basic condition for this system to be central. Some of the images used to perform the calibration of this system under the different models are shown in Fig. 9. As in the hyper-catadioptric case we compute the mirror parameter and the principal point. This result is shown in Table 6. We observe that DLT-like and Sphere-2DPattern give similar mirror hyperbolic parameter and Sphere-Lines estimates.
<table>
<thead>
<tr>
<th></th>
<th>3D Error (mm)</th>
<th>Reprojection Error (pixels)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>σ</td>
</tr>
<tr>
<td>Sphere-2DPattern</td>
<td>4.10</td>
<td>2.40</td>
</tr>
<tr>
<td>DLT-like</td>
<td>6.44</td>
<td>4.85</td>
</tr>
<tr>
<td>Sphere-Lines</td>
<td>12.42</td>
<td>7.68</td>
</tr>
<tr>
<td>Distortion-2DPattern</td>
<td>2.61</td>
<td>1.83</td>
</tr>
</tbody>
</table>

Table 5: Reprojection error in pixels from two images used to compute the SfM using the unknown-shape mirror.

(a) (b)

(c) (d)

Figure 9: Some images used to calibrate the non-central catadioptric system. (a) Sphere-Lines. (b) DLT-like approach. (c,d) Sphere-2DPattern and Distortion-2DPattern approaches.

a parabolic mirror with $\xi = 1$. The 3D error and the reprojection error are shown in Table 7. We observe again that all the approaches have a similar behavior even with a non-central system. We observe that the reprojection error for all the methods is below one pixel. The Sphere-2DPattern gives the smallest reconstruction error (0.96mm). The maximum error is given by the Sphere-Lines approach with an error of 2.44 mm. The other two methods (DLT-like, Distortion-2DPattern) give similar errors close to 1.50 mm.

5.6. Discussion

After all these experiments with different systems we observe that all the approaches give similar results, with the exception of the Sphere-Lines approach with the fish-eye system. We consider here the importance on what we need to make these methods to work and the problems observed at the calibration time.

- **Sphere-2DPattern, Distortion-2DPattern.** These two approaches require multiple images of a 2D pattern to perform the calibration. Both of them
Figure 10: Images used in the SfM experiment with reprojected points superimposed (non-central hyper-catadioptric).

<table>
<thead>
<tr>
<th>Method</th>
<th>Error (mm)</th>
<th>Reprojection Error (pixels)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>σ</td>
</tr>
<tr>
<td>Sphere-2DPattern</td>
<td>0.96</td>
<td>0.45</td>
</tr>
<tr>
<td>DLT-like</td>
<td>1.63</td>
<td>0.87</td>
</tr>
<tr>
<td>Sphere-Lines</td>
<td>2.44</td>
<td>1.35</td>
</tr>
<tr>
<td>Distortion-2DPattern</td>
<td>1.38</td>
<td>0.77</td>
</tr>
</tbody>
</table>

Table 6: Comparison of the physical parameters given by the 3 methods based on the sphere model in the non-central system.

have automatic corner extractors but most of the times these do not work properly and the points have to be given manually. This is the most tedious part since we have a minimum of eight to ten images, each image containing 48 points giving a total of 384 \(\sim 480\) points. Besides that the Sphere-2DPattern approach requires the user to indicate the image center and a minimum of three non-radial points to estimate the focal length.

- **DLT-like.** This approach does not require any prior information but one single omnidirectional image containing 3D points spread on three different planes. The inconvenient with this method is to obtain the 3D points contained in the 3D pattern. This is performed with several images of the 3D pattern and a photogrammetric software. All the image points in the 3D pattern images are given manually. We observed that depending on the image used the mirror parameter \(\xi\) is better or worse estimated. Something similar happens with the Sphere-2DPattern approach.

- **Sphere-Lines.** This approach requires the easiest setting to be constructed.
It only requires one single omnidirectional image containing a minimum of 3 lines. One thing observed using this approach is that it hardly depends on the principal point estimation. If this estimation is not accurate enough the calibration is not performed correctly. Also we observe some difficulties while calibrating the unknown shape catadioptric system. The number and the location of lines in the image is important to correctly calibrate the system. Many times using more than three lines we obtained calibrations containing non-real solutions. We also observe with this catadioptric system some convergence problems using three and more than three lines.

One common thing to all these approaches is that either lines or points have to cover as much as possible the omnidirectional image. Particularly on the periphery area. All these approaches perform a non-linear step after the initialization of the intrinsic and extrinsic parameters is computed. This step is quite important. As it is commented in [11], the initialization of the mirror parameter does not have a strong influence because the non-linear step will reach the right values.

6. Conclusions

In this paper we have presented the comparison of four methods to calibrate omnidirectional cameras available as OpenSource. Two of them require images of a 2D pattern, one requires images of lines and the last one requires one image of a 3D pattern. Three of these approaches use the sphere camera model. This model can give some information about the mirror present in the catadioptric system besides it provides a theoretical projection function. The other approach is based on a distortion function. Both models can deal with any central catadioptric system and fish-eye. However the Sphere-Lines approach that uses the sphere camera model cannot deal with the fish-eye system. All these approaches use a non-linear step which allows them to have a reprojection error less than 1 pixel. In this paper we perform a SfM experiment to compare the different approaches with a useful criteria. This experiment showed that the calibration reached by any of these methods can be give accurate reconstruction results. The distribution of the points in the omnidirectional images is important in order to have an accurate calibration. These points have to cover as much as possible in the omnidirectional image and mainly in the peripheric area.

References


