HAL
open science

Calibration of omnidirectional cameras in practice: A comparison of methods

Luis Puig, Jesús Bermúdez, Peter Sturm, Josechu Guerrero

To cite this version:

Luis Puig, Jesús Bermúdez, Peter Sturm, Josechu Guerrero. Calibration of omnidirectional cameras in practice: A comparison of methods. Computer Vision and Image Understanding, 2012, Special Issue: Virtual Representations and Modeling of Large-scale Environments (VRML), 116 (1), pp.120-137. 10.1016/j.cviu.2011.08.003 . hal-00644989

HAL Id: hal-00644989

https://inria.hal.science/hal-00644989

Submitted on 28 Nov 2011

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

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

Calibration of Omnidirectional Cameras in Practice. A Comparison of Methods

Luis Puig[*,a], J. Bermúdez[b], Peter Sturm[c], J. J. Guerrero[a]

[a] *Instituto de Investigación en Ingeniería de Aragón-Departamento de Informática e Ingeniería de Sistemas, Universidad de Zaragoza, Spain*
[b] *Instituto Tecnológico de Aragón ITA, Zaragoza, Spain*
[c] *INRIA Rhône-Alpes and Laboratoire Jean Kuntzmann, Grenoble, France*

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 camera models and often limited to 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 comparison using simulated and real data. We choose methods available as OpenSource and which do not require a complex pattern or scene. The evaluation protocol 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 has widely increased. The major advantage of this type of cameras is their wide field of view (FOV) which allows them to have a view of the whole scene. They have been used in such different areas as surveillance, tracking, visual navigation, localization and SLAM, structure from motion, active vision, visual odometry, photogrammetry, camera networks, reconstruction of cultural heritage, etc. There exist several types of omnidirectional cameras which can be classified as central and non-central. Among the non-central cameras we can find the rotating camera, which consists of a conventional camera with a mechanic system that allows it to move along a circular trajectory and to acquire images from the surroundings. Polycameras which are camera clusters of conventional cameras pointing to different directions in a particular configuration. Another type of non-central systems are dioptric systems which use wide-angle lenses such as fish-eye lenses combined with conventional cameras. The central omnidirectional cameras are those which satisfy the single-viewpoint property. This is an important property since it allows to easily calculate the directions of light rays coming into the camera. Baker and Nayar [1] extensively studied the catadioptric systems, combinations of camera lenses and mirrors. They proved that the elliptic, parabolic and hyperbolic mirrors, combined with conventional cameras, are the only ones that ensure the single-viewpoint property, provided that the mirror is positioned appropriately relative to the camera. The two most popular of such systems are the hyper-catadioptric and the para-catadioptric ones. The former is composed by a hyperbolic mirror and a perspective camera. The latter is composed by a parabolic mirror and an orthographic camera. In combination with the fish-eye lenses these are the three omnidirectional system most used by the computer vision and robotics communities.

*Corresponding author
Email addresses: lpuig@unizar.es (Luis Puig), jesusbermudezcameo@gmail.com (J. Bermúdez), Peter.Sturm@inrialpes.fr (Peter Sturm), jguerrer@unizar.es (J. J. Guerrero)

There exist several geometric and analytic models to deal with omnidirectional systems, see Sturm et al. [2]. In the case of central catadioptric systems Svoboda and Pajdla [3] propose different models for different mirrors and give formulae for the associated epipolar geometry. Strelow et al. [4] deal directly with the reflection properties of the rays on the mirror. A unified model was proposed by Geyer and Daniilidis [5]. 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 [6] and Ying and Hu [7]. This sphere model is the most used model in current days. With respect to the non-central systems, in particular for fish-eye lenses we can find the following approaches. Swaminathan and Nayar [8] model this type of projections as a combination of three types of distortion. These distortions are the shift of the optical center, radial distortion and decentering distortion. Micusik and Pajdla [9] compute the projection of 3D points to the camera plane using trigonometric functions which are linearized through Taylor series. This is done for a particular type of camera. Kannala and Brandt [10] propose a generic model to deal with all cameras equipped with fish-eye lenses. They consider the projections as a series of odd powers of the angle between the optical axis and the incoming ray, then they complete the model by adding radial and tangential distortion. Courbon et al. [11] propose a generic model to calibrate any fish-eye system based on the sphere camera model. Another category that should be mentioned is the generic methods that can model any arbitrary imaging system. Grossberg and Nayar [12] 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. [13].

Most of the applications mentioned before require to recover metric information from the environment. This 3D information is crucial when the omnidirectional cameras interact in real scenarios. The metric information depends entirely on the complete calibration of the omnidirectional system. For these practical applications the camera calibration is a basic step for subsequent and higher level tasks. The accuracy of such applications relies on the accuracy of the camera calibration. A considerable number of approaches to either calibrate central catadioptric systems or to calibrate fish-eye lens systems or both have been developed. Moreover, as we have observed they can use different projection models. With respect to central catadioptric systems, there exist some approaches that separate the calibration of the perspective/orthographic camera from the computation of the mirror parameters [3, 14]. However most of them deal with the catadioptric system as a whole. Some of these methods use single or multiple views of a 2D pattern [10, 15, 16, 17, 18, 19], 3D pattern [20] previously presented in [21], cylinder pattern [22], some others use a single image containing features like lines [23, 24, 25, 26, 7, 27, 28, 29, 30, 31]. Finally there are other methods that perform a self-calibration of the system [32, 33, 34, 35].

Contributions and Related Work

As observed above there exist many calibration methods. They use different techniques and models to calibrate the omnidirectional systems. Some works have tried either to classify or to compare them. Ying and Hu [7] classify the calibration methods in three categories: i) Known World Coordinates; ii) Self-calibration and iii) Projection of lines. They only consider 7 methods. On the other hand, they do not perform any comparison with any of such methods. A more specific classification is given by Deng et al. [17] where 10 methods are grouped in five categories: i) Self-calibration; ii) Sphere-based calibration; iii) Line-based calibration; iv) Point-based calibration and v) 2D calibration. We observe that methods based on 2D patterns have appeared emulating calibration methods for conventional cameras. The sphere-based category only contains one method, which also uses lines and it could be classified in that category. In that work there is no comparison of the proposed method to any other. Frank et al. [18] identify three big groups of calibration methods: i) Known World Coordinates which include those based on 2D patterns, which from our point of view should belong to different categories; ii) Geometric Invariants which include the methods based on projections of lines and iii) Self-calibration methods. They mention a total of 8 methods. A comparison of the proposed method with the online available methods [15] and [16] is presented. They use 4 different data sets including two fisheye, a paracatadioptric system and a system with very small distortion. Since method [16] does not allow the manual extraction of grid points, the authors only consider those images where the grid points are extracted successfully, having a limited set of images. This situation has as consequence a poor calibration of the system. In our paper we extract the grid points manually,

allowing the methods to have the maximum data available, which permits to obtain the best performance and in consequence to perform a fair comparison of the methods. Toepfer and Ehlgen [22] do not present a classification but a comparison of their proposed method with [15, 16, 36]. The performance of the methods is given considering a combination of the root mean square error (we assume the reprojection error) with the number of parameters of the method, through the principle of minimum description length. However it is not clear which catadioptric system has been calibrated neither how the method [36] is adapted to work with catadioptric systems.

As a contribution of this paper we present a classification of the existing approaches to calibrate omnidirectional systems. We propose five categories based on the main entity required by each method to perform the calibration of the systems. We also present in Table 1 the relevant information of each method according to valuable criteria: pattern or entity required, considering the minimum number of elements, number of views, analytical model, type of mirror, or if the method requires the separate calibration of the mirror and the camera.

Since the amount of calibration methods is considerable, the selection of a particular calibration method seems to be difficult and even more so if we consider the problems involved in the implementation process. Among all approaches mentioned previously, there is a group of calibration methods for omnidirectional systems (catadioptric and fish-eye) available online as OpenSource toolboxes. These methods can save time and effort when the goal is beyond the calibration itself and when the user is more interested in obtaining 3D motion and structure results than to deal with complex projection models. In this work, we evaluate these methods and provide an analysis with simulations and real images. Moreover, we use a Structure from Motion application with two omnidirectional images, where we become users of the calibrations provided by these approaches. This experiment shows the behavior of the approaches in a real scenario. Besides the performance, we also consider the ergonomics and ease of usage, as well as the type of features, the algorithm and the type of pattern, since they are important elements that can help the user to select the best approach. Moreover we present an up-to-date list of the calibration methods developed that consider catadioptric and fish-eye systems, allowing the reader to decide to implement a different method. These calibration methods are:

1. Mei and Rives [16][1] which uses the sphere camera model and requires several images of a 2D pattern. We will call this approach *Sphere-2DPattern*.
2. Puig et al. [20][2] which obtains a solution in closed form requiring a set of 3D-2D correspondences. It also uses the sphere camera model. We call this approach *DLT-like*.
3. Barreto and Araujo [25][3] 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. [15][4] 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. In section 2 we present a review of the existing calibration methods up to now, doing a classification based on the main entity required by the approaches to calibrate the omnidirectional systems. 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. Classification of the 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 after that

[1] `http://www.robots.ox.ac.uk/~cmei/Toolbox.html`
[2] `http://webdiis.unizar.es/~lpuig/DLTOmniCalibration`
[3] `http://www.isr.uc.pt/~jpbar/CatPack/main.htm`
[4] `http://asl.epfl.ch/~scaramuz/research/Davide_Scaramuzza_files/Research/OcamCalib_Tutorial.htm`

find the mirror parameters. In this section we present a review and classification of existing calibration methods. We propose five categories based on the main entity required to perform the calibration of the systems.

- **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 [23] calibrate para-catadioptric cameras from the images of only three lines. Barreto and Araujo [26] 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 [7] analyze the relation of the camera intrinsic parameters and imaged sphere contours. 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 [27] 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 image detection which ensures that all detected conics must belong to a line image family related to certain intrinsic parameters. Vandeportaele et al. [28] slightly improves [23] 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 a unified manner. Wu et al. [29] 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. [30] calibrate a system where the perspective camera is placed in a generic position with respect to a mirror, i.e., a non-central system. They 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. [31] 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 [37] 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. [15] 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 [16] 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 additional radial and tangential distortion parameters are added to consider real-world errors. Deng et al. [17] 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. Gasparini et al. [19] compute the image of the absolute conic (IAC) from at least 3 homographies which are computed from images of planar grids. The intrinsic parameters of the central catadioptric systems are recovered from the IAC.

- **3D Point-based calibration**. These methods require the position of 3D points observed usually in a single image. Aliaga [38] proposes 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 [39] introduced the invariants of 1D/2D/3D space points and use them to compute the camera principal point with a quasi-linear method. Puig et al. [20] present an approach based on the Direct Linear Transformation (DLT) using lifted coordinates to calibrate any central catadioptric camera. 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 [32] uses the consistency of pairwise tracked point features for calibration. The method is only suitable for para-catadioptric systems. Micusik and Pajdla [33] 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. Ramalingam et al. [34] use pure translations and rotations and the image matches to calibrate central cameras from geometric constraints on the projection rays. Espuny and Burgos Gil [35] developed a similar approach that uses two dense rotational flows produced by rotations of the camera about two unknown linearly independent axes which pass through the projection center.

- **Polarization Imaging**. This method is proposed by Morel and Fofi [14]. 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.

3. Sphere Camera model

As many of the calibration methods are based on the sphere camera model [5, 6], 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, $\mathbf{r}_\pm$. These points are then projected into a normalized plane π_{nor} resulting in two points, $\mathbf{q}_\pm$. Finally these points are projected into the perspective plane π_p giving again two points $\mathbf{x}_\pm$, 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 and ψ the distance between the image plane and the projection center. $\xi = 0$ 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 $\mathbf{C}_p = (0, 0, -\xi)^\mathsf{T}$. The perspective camera is modeled by the projection matrix $\mathsf{P} \sim \mathsf{A}_p\mathsf{R}_p \begin{pmatrix} \mathtt{I} & -\mathbf{C}_p \end{pmatrix}$, 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 $\mathsf{A}_p\mathsf{R}_p$ 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 (ξ, ψ) which are computed from the mirror parameters [26, 16]. The **intrinsic parameters of the catadioptric camera** are thus ξ and H_c, where H_c is defined as

$$\mathsf{H}_c = \mathsf{A}_p\mathsf{R}_p \underbrace{\begin{bmatrix} \psi - \xi & 0 & 0 \\ 0 & \xi - \psi & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathsf{M}_c} \tag{1}$$

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 and Rives [16] just take into account the difference $\xi - \psi$ and call it η, so the flip is not considered in that model. Following that, (2) encodes this two projections $(\mathbf{x}_\pm)$ and its corresponding inverse function is (3), which maps an image point $\mathbf{x}$ to an oriented 3D ray $\mathbf{X}$.

Method	Pattern/Entity	Number of views	Model	Mirror	W/S
Kang [32]	10 point tracks	Multiple	Particular Central Catadioptric	Parabolic	W
Svoboda and Pajdla [3]	—	Single	Particular Central Catadioptric	Generic	S
Caglioti et al. [30]	1 line + mirror contours	Single	Part. Non-central Catadioptric	Generic	W
Aliaga [38]	3D known points	Multiple	Part. Non-central Catadioptric	Parabolic	W
Toepfer and Ehlgen [22]	3D pattern/2D pattern (multiple points)	Single/ Multiple	Particular Central Catadioptric	Hyperbolic & Parabolic	W
Geyer and Daniilidis [23]	3 lines	Single	Sphere	Parabolic	W
Geyer and Daniilidis [24]	2 vanishing points	Single	Sphere	Parabolic	W
Barreto and Araujo [26]	3 lines	Single	Sphere	Generic	W
Ying and Hu [7]	2 lines/3 spheres	Single	Sphere	Generic	W
Ying and Zha [27]	3 lines	Single	Sphere	Generic	W
Vandeportaele et al. [28]	3 lines	Single	Sphere	Parabolic	W
Wu et al. [29]	lines	Single	Sphere	Parabolic	W
Wu et al. [31]	3 lines	Single	Sphere	Generic	W
Vasseur and Mouaddib [37]	lines (minimum no. n/a)	Single	Sphere	Generic	W
Mei and Rives [16]	2D pattern (multiple points)	Multiple	Sphere	Generic	W
Puig et al. [20]	3D pattern (20 3D-2D correspondences)	Single	Sphere	Generic	W
Deng et al. [17]	2D pattern (multiple points)	Multiple	Sphere	Generic	W
Gasparini et al. [19]	2D pattern (multiple points)	Multiple	Sphere	Generic	W
Wu and Hu [39]	4 correspondences	Multiple	Sphere	Generic	W
Scaramuzza et al. [15]	2D pattern (multiple points)	Multiple	Distortion	Generic	W
Frank et al. [18]	2D pattern (multiple points)	Multiple	Distortion	Generic	W
Micusik and Pajdla [33]	9 correspondences (epipolar geometry)	Multiple	Distortion	Generic	W
Morel and Fofi [14]	3 polarized images	Multiple	Generic Camera*	Generic	S
Ramalingam et al. [34]	2 Rotational & Translational Flows	Multiple	Generic Camera*	Generic	W
Espuny and Burgos Gil [35]	2 Rotational Flows	Multiple	Generic Camera*	Generic	W

Table 1: Classification of the calibration methods for omnidirectional systems. W = whole system. S = separate calibration (1. camera and 2. mirror parameters). *They use the same approach [40].

$$\mathbf{x}_{\pm} = \hbar(\mathbf{X}) = \begin{bmatrix} X \\ Y \\ Z \pm \xi\sqrt{X^2+Y^2+Z^2} \end{bmatrix} \tag{2}$$

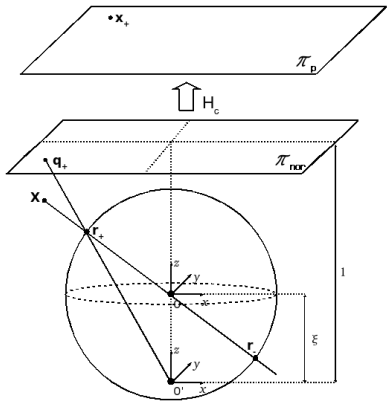


Figure 1: Projection of a 3D point in the image plane using the sphere camera model.

$$\hbar^{-1}(\mathbf{x}) = \begin{bmatrix} \frac{\xi \mp \sqrt{1+(1-\xi^2)(x^2+y^2)}}{x^2+y^2+1} x \\ \frac{\xi \mp \sqrt{1+(1-\xi^2)(x^2+y^2)}}{x^2+y^2+1} y \\ \frac{\xi \mp \sqrt{1+(1-\xi^2)(x^2+y^2)}}{x^2+y^2+1} \mp \xi \end{bmatrix} \tag{3}$$

Some authors [26, 16] just take into account the positive projection given by function $\hbar$ but some others [41, 42] use lifted coordinates to deal with the non-linearities present in function $\hbar$. The vector lifting consists in mapping a 3-vector $\mathbf{x} = (x_1, x_2, x_3)$ in the projective space $\wp^2$ to a 6-vector $\hat{\mathbf{x}}$ in the projective space $\wp^5$ and performed by the following equation

$$\hat{\mathbf{x}} = \begin{pmatrix} x_1^2 & x_1x_2 & x_2^2 & x_1x_3 & x_2x_3 & x_3^2 \end{pmatrix}^\mathsf{T} \tag{4}$$

Lifted matrices are also used to compute the generic catadioptric projection matrix. Two different ways to compute this lifting can be found in [21, 41].

4. Calibration Methods analyzed

In this section we summarize the four OpenSource methods used to calibrate omnidirectional systems. The purpose of this section is to show a general view of the methods that help the reader to understand the core of each method.

4.1. Sphere-2DPattern approach

This approach [16] 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 many 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.

This approach uses a total of 17 parameters to relate a scene point to its projection in the image plane:

- Seven extrinsic parameters $(\mathbf{q}, \mathbf{t})$ representing the relation between the camera reference system and the world reference system. A 4-vector $\mathbf{q}$ represents the rotation as a quaternion and a 3-vector $\mathbf{t}$ represents the translation.

- One mirror parameter ξ.
- Four distortion parameters $Dist = [k_1, k_2, p_1, p_2]$, two for tangential distortion and two for radial distortion [43].
- Five parameters representing a generalized camera projection $P = [\theta, \gamma_1, \gamma_2, u_0, v_0]$. (γ_1, γ_2) are the focal lengths of the catadioptric system for x and y axis, θ is the skew parameter, and (u_0, v_0) is the principal point.

The 2D pattern used to calibrate the camera is composed of m points $\mathbf{X}_i$ with their associated image values $\mathbf{x}_i$. The solution of the calibration problem is obtained by minimizing the reprojection error using the Levenberg-Marquardt algorithm.

4.2. Sphere-Lines approach

This method [26] is based on computing the absolute conic $\hat{\Omega}_\infty = \mathsf{H}_c{}^{-\mathsf{T}}\mathsf{H}_c{}^{-1}$ and the mirror parameter ξ 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 mirror parameter ξ is computed. From the image of a conic in the omnidirectional images it 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 $\hat{\Omega}_\infty$. Once the image of the absolute conic $\hat{\Omega}_\infty$ is computed, from its Cholesky decomposition we obtain

$$\mathsf{H}_c = \begin{pmatrix} \gamma_x & \beta & c_x \\ 0 & \gamma_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \tag{5}$$

with the intrinsic parameters γ_x and γ_y (focal lengths), skew (β) and principal point (c_x, c_y).

4.3. DLT-like approach

This approach [20] 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 6×10 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.

In this approach a 3D point $\mathbf{X}$ is mathematically projected to two image points $\mathbf{x}_+$, $\mathbf{x}_-$, which are represented in a single entity via a degenerate dual conic Ω. The relation between them is $\Omega \sim \mathbf{x}_+\mathbf{x}_-^\mathsf{T} + \mathbf{x}_-\mathbf{x}_+^\mathsf{T}$.

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

$$\mathbf{c} \sim \widehat{\mathsf{H}}_{c6\times6}\mathsf{X}_\xi\widehat{\mathsf{R}}_{6\times6}\begin{pmatrix}\mathtt{I}_6 & \mathsf{T}_{6\times4}\end{pmatrix}\widehat{\mathbf{X}}_{10} \tag{6}$$

where, R represents the rotation of the catadioptric camera. X_ξ a 6×6 matrix and $\mathsf{T}_{6\times4}$ depend only on the sphere model parameter ξ and position of the catadioptric camera $\mathbf{C} = (t_x, t_y, t_z)$ respectively. Thus, a 6×10 catadioptric projection matrix, P_{cata}, is expressed by its intrinsic A_{cata} and extrinsic T_{cata} parameters

$$\mathsf{P}_{cata} = \underbrace{\widehat{\mathsf{H}}_c\mathsf{X}_\xi}_{\mathsf{A}_{cata}}\underbrace{\widehat{\mathsf{R}}_{6\times6}\begin{pmatrix}\mathtt{I}_6 & \mathsf{T}_{6\times4}\end{pmatrix}}_{\mathsf{T}_{cata}} \tag{7}$$

This matrix is computed from a minimum of 20 3D-2D lifted correspondences in a similar way to the perspective case [44] using least squares

$$\left(\widehat{[\mathbf{x}]}_\times \otimes \widehat{\mathbf{X}}\right)\mathbf{p}_{cata} = \mathbf{0}_6 \tag{8}$$

The 60-vector $\mathbf{p}_{cata}$ contains the 60 coefficients of P_{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 uses 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 [15] the projection model is different from the one previously presented. The only assumption is that the image projection function can be described by a polynomial, based on 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 highly distorted image and we have to compute the distortion parameters to obtain such a distorted image. This approach, like Sphere-2DPattern, requires several images from different unknown positions of a 2D pattern. The accuracy depends on the number of images used and on the degree of the polynomial.

Under this model a point in the camera plane $\mathbf{x}' = [x', y']^{\mathsf{T}}$, is related to a vector $\mathbf{p}$ which represents a ray emanating from the viewpoint O (located at the focus of the mirror) to the scene point $\mathbf{X}$. This relation is encoded in the function $\mathbf{g}$

$$\mathbf{p} = \mathbf{g}(x') = \mathsf{P}\mathbf{X} \tag{9}$$

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

$$\mathbf{g}(\mathbf{x}') = (x', y', f(x', y'))^{\mathsf{T}}, \tag{10}$$

and f is defined as

$$f(x', y') = a_0 + a_1\rho' + a_2\rho'^2 + \cdots + a_n\rho'^n \tag{11}$$

where ρ' is the Euclidean distance between the image center and the point. 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.

The camera calibration under this approach is performed in two steps. The first step computes 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. The authors do not mention it, but after this linear process a non-linear optimization is performed using the Gauss-Newton algorithm[5].

5. Experiments

In order to compare the different calibration methods we calibrate a hyper-catadioptric system[6], a fish-eye and a commercial proprietary shape mirror[7], which we name unknown-shape system. Additionally, we displace the perspective camera of the hyper-catadioptric system far from the mirror. This allows the displacement of the optical center of the perspective camera from the other focus described by the hyperbolic mirror, leading to a non-central system. We calibrate these four systems with the four methods and compare

[5]This algorithm is provided by the **lsqnonlin** Matlab function
[6]Neovision H3S with XCD-X710 SONY camera
[7]http://www.0-360.com

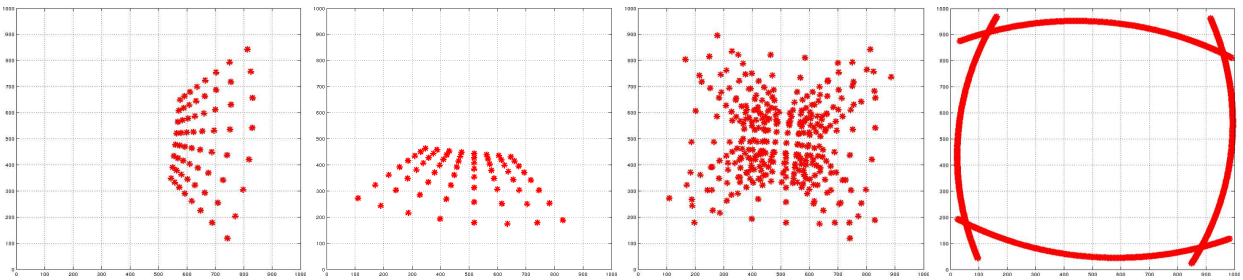


Figure 2: Synthetical images used by the calibration methods. Sphere-2DPattern and Distortion-2DPattern approaches use five images similar to (a) and (b). DLT-like approach uses a single image containing 5 planar patterns (c) and Sphere-Lines approach use 4 image lines (d).

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 outer mirror border to compute the principal point. Then it asks for four aligned edge points on a non-radial line to compute the focal length. With this information it asks for the four corner points of the pattern and uses a subpixel technique to extract the rest of the points present in the pattern. If the focal length 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 at least three different planes with known relative position.

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.

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.

5.1. Evaluation Criteria

Previous to the comparison of the real system we perform an experiment using simulated data. The purpose of this experiment is to observe the behavior of all approaches under optimal conditions and to measure their sensitivity to noise. We simulate two central catadioptric systems, a hypercatadioptric[8] with mirror parameter $\xi = 0.7054$ and a paracatadioptric system. Firstly, we generate five synthetic images, each one containing a calibration pattern, that cover the whole FOV of the omnidirectional image, two of these images are shown in Fig. 2(a) and Fig. 2(b). These images are used by the Sphere-2DPattern and the Distortion-2D approaches. We combine the five calibration patterns in a single image (Fig. 2(c)) that is used by the DLT-Like approach. These three approaches use the same points. In the case of the Sphere-Lines approach four lines are present in a single image with a total of 904 points (Fig. 2(d)). We add Gaussian noise to the image coordinates. For every noise level σ (in pixels) we repeat the experiment 100 times in order to avoid particular cases due to random noise. In Fig. 3 and Fig. 4 we show the mean and the standard deviation of the reprojection error corresponding to the hypercatadioptric and the paracatadioptric systems, respectively for the analyzed approaches.

As we observe the behavior of the four approaches is quite similar. All of them respond in the same way to the amount of noise present in the images. This experiment shows that under optimal circumstances the performance of the four approaches is quite similar. One thing we observe with respect to the Sphere-Lines

[8] http://www.accowle.com

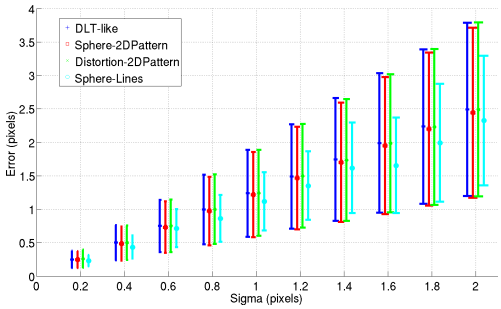


Figure 3: Reprojection error in pixels as a function of noise (σ) corresponding to the hypercatadioptric system.

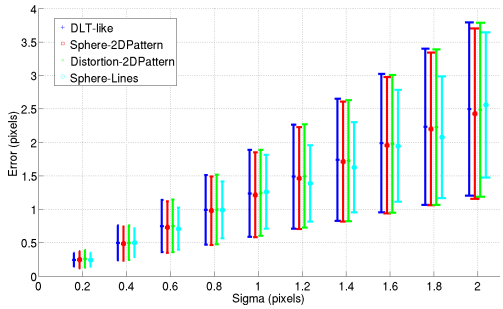
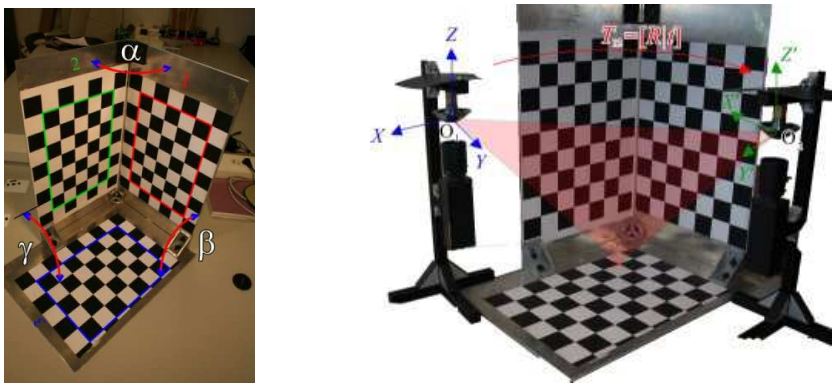


Figure 4: Reprojection error in pixels as a function of noise (σ) corresponding to the paracatadioptric system.



(a) (b)

Figure 5: (a) 3D pattern. (b) SfM configuration to test the calibrations.

approach is that four lines are enough to compute the calibration. We try to calibrate the systems by using a higher number of lines. This caused the approach to become slower and in some occasions not to converge at all. This behavior is due to the lines containing a high number of points, therefore increasing 2 or more lines means to increase hundreds of points.

We observe that the reprojection error is not sufficient to decide which method is endowed with the best performance, however it is a necessary condition to qualify an algorithm as performing well. Moreover, a disadvantage of this error measure is that we can make it smaller by just adding more parameters to a model. As an example we can observe the reprojection error of the hypercatadioptric system given by the Sphere-2DPattern approach (see Table 3). The obtained reprojection error during calibration was considerably small, only 0.000005 pixels, which could be considered as zero. This approach adds five distortion parameters: three for radial distortion and two for tangential distortion. Originally, these parameters are not considered in the model. To verify the impact of these parameters in the calibration we repeated the calibration. When we only consider radial distortion (3 parameters), the reprojection error increased to 0.42 pixels. When no distortion is considered at all, the reprojection error increased to 0.64 pixels. As previously stated, the reprojection error is not definitive in showing which approach performs the best.

Instead of the reprojection error we choose to consider a Structure from Motion task from two calibrated omnidirectional images of a 3D pattern (Fig. 5(a)) built in our lab. The pattern has been measured with high accuracy using photogrammetric software[9]. Thus, a 3D reconstruction by bundle adjustment has been made. We used 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. 5(b) shows

[9]PhotoModeler software was used.

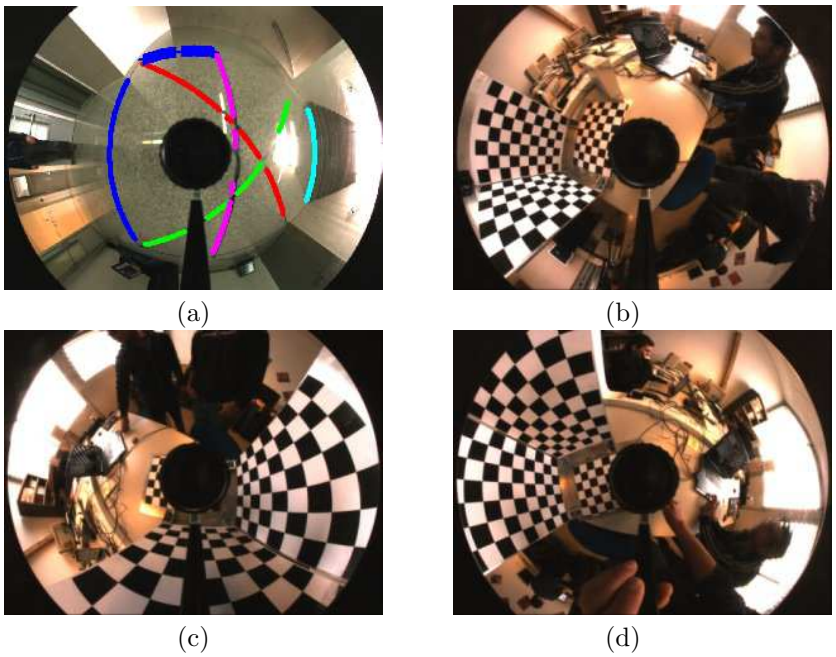


Figure 6: Some images used to calibrate the hyper-catadioptric system.(a) Sphere-Lines. (b) DLT-like approach. (c,d) Sphere-2DPattern and Distortion-2DPattern approaches.

the configuration used for the SfM experiment. Using the calibration provided by each method we compute the corresponding 3D rays from each omnidirectional image. The correspondences between the two images were given manually. We use these correspondences to compute the essential matrix E which relates them. From this matrix we compute two projection matrices $\mathsf{P}_1 = [\mathsf{I}|\mathbf{0}]$ and $\mathsf{P}_2 = [\mathsf{R}|\mathbf{t}]$. Then, with these projection matrices and the 3D rays we compute an initial 3D reconstruction using a linear triangulation method [44] 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. 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. 6 we observe some of the images used to calibrate this system. We use one image to calibrate the system using the Sphere-Lines and the DLT-like approaches. Eleven images of the 2D pattern were used to calibrate the system using both 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 ξ since they are related with sphere model parameters. The Distortion-2DPattern approach does not offer information about the catadioptric system. As we can see,

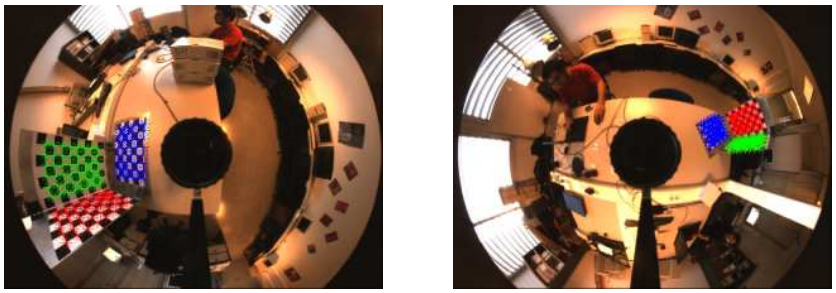


Figure 7: Images used in the SfM experiment with reprojected points superimposed (hyper-catadioptric).

	ξ	(u_0, v_0)
Manufacturer Info	0.9662	(511.88,399.25)
Sphere-2DPattern	0.9684	(513.93,400.76)
Sphere-Lines	1.0215	(523.82, 416.29)
DLT-like	0.9868	(509.95, 398.54)

Table 2: Comparison of the physical parameters given by the 3 methods based on the sphere model.

Method	Original paper	Calibration	Structure from Motion
Sphere-2DPattern	0.40	0.00005	0.34
Sphere-Lines	n/a	1.11*	1.19
DLT-Like	0.30	0.47	0.40
Distortion-2DPattern	1.2	0.82	0.45

Table 3: Reprojection error from different sources corresponding to the hypercatadioptric system. *since the Sphere-Lines does not use the reprojection error we take it from simulations with Gaussian noise $\sigma = 1 pixel$.

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 the Sphere-2DPattern method asks the user to give the image center and the rim of the mirror to compute the principal point and the DLT-like algorithm does not need any of this a priori information but requires known positions of the 3 planes in the pattern.

In Fig. 8 we show the histogram corresponding to the accuracy in millimeters of the 144 reconstructed 3D points. We observe that the Sphere-2DPattern approach has the highest number of reconstructed points closer to the ground truth. The highest error corresponds to Sphere-Lines and Distortion-2DPattern with one point 5 mm far from the ground truth. In Fig. 9 we show the reprojection error of the 288 points of the two images used in the experiment. We observe that all the approaches are below the 2 pixel error and three of them within the 1 pixel error.

The hypercatadioptric system is the more complex catadioptric system under the sphere camera model since the mirror parameter ξ is in the interval $]0, 1[$. In opposition, the paracatadioptric system where $\xi = 1$ simplifies considerably the model. In this context we decide to present the reprojection error corresponding to this system from three different sources (see Table 3): the reprojection error shown in the original paper where the approach was firstly presented (first column), the reprojection error obtained when we calibrate the system, and the reprojection error from the Structure from Motion experiment. This information provides the reader with a clear idea about the performance of all approaches under different circumstances and allows us to observe that the reprojection error given at the calibration time is less informative that the reprojection error of a Structure from Motion experiment where the calibration is an early step and all the approaches are under the same circumstances.

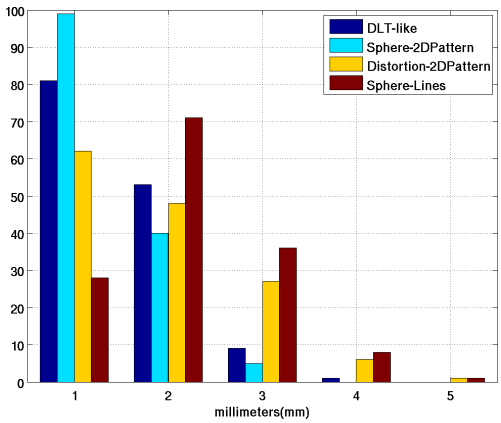


Figure 8: Number of reconstructed 3D points within an error distance in millimeters using a hyper-catadioptric system.

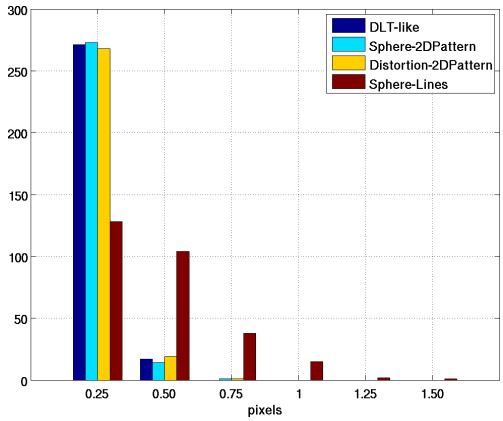


Figure 9: Number of reprojected points within an error distance in pixels using a hyper-catadioptric system.

5.3. Fish-eye Lens

The fish-eye lens used in this experiment is a Raynox DCR-CF185PRO with a FOV of 185° on all directions. It is mounted on a high-resolution camera. In Fig. 10 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 and only one image using the DLT-like approach. We use the image of 7 lines to perform the calibration with the Sphere-Lines approach. Since none of the methods provides any information about the system we just show the results obtained from the SfM experiment.

In Fig. 12 and Fig. 13 we show the results of this experiment. We observe that the methods that claim to be able to calibrate the non-central fish eye systems indeed 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 this camera. It gives reprojection errors close to one hundred pixels. We observe that the three valid approaches give similar results. The DLT-like gives the best results with the highest number of reconstructed points within 1 mm error, although it is not designed to handle fish-eye lenses. With respect to the reprojection error, we observe a similar behavior of the three approaches with a maximum error of 2 pixels.

5.4. Unknown-shape Catadioptric System

This system is the combination of a commercial proprietary shape mirror and a high-resolution camera. 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 results contained complex numbers or there were problems of convergence. In Fig. 14 we observe some of the images used to calibrate this system under all the analyzed approaches. In Fig. 15 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. 15 (a), (b)). The results of the experiments, the 3D accuracy and the reprojection error are shown in Fig. 16 and Fig. 17, respectively. We observe that the Distortion-2DPattern approach has the highest number of points within the 5 mm error, 131 out of 144. The worst performance is given by the Sphere-Lines approach with maximum errors of 35 mm. The other two approaches show similar performance, with the majority of the reconstructed 3D points within the 10 mm error. The lowest reprojection error is given by the Distortion-2DPattern with all the points within the 1 pixel error and the other three approaches have a similar behavior with the majority of the reprojected points within the 3 pixel error.

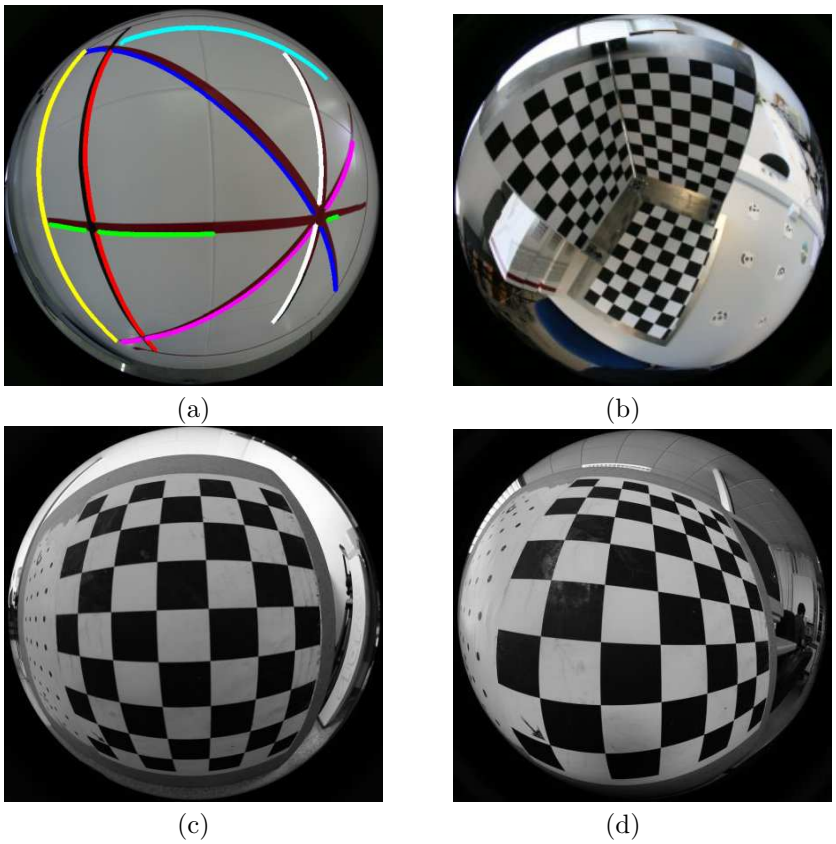


Figure 10: Some images used to calibrate the fish-eye system.(a) Sphere-Lines. (b) DLT-like approach. (c,d) Sphere-2DPattern and Distortion-2DPattern approaches.

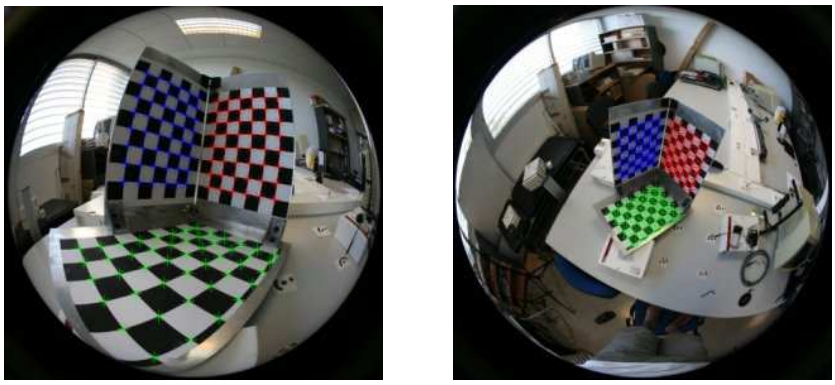
5.5. Non-central hyper-catadioptric system

This system is the one described in the hyper-catadioptric experiment. The only difference is that the perspective camera is displaced as far as possible from the mirror. This causes that the optical center of the perspective camera is not located at the other focus of the hyperbolic mirror, which is 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. 18. As in the hyper-catadioptric case we compute the mirror parameter and the principal point. This result is shown in Table 4. We observe that DLT-like and Sphere-2DPattern give similar mirror hyperbolic parameter and Sphere-Lines estimates a parabolic mirror with $\xi = 1$.

The 3D error and the reprojection error are shown in Fig. 20 and Fig. 21, respectively. We observe that all the approaches have a similar behavior even with a non-central system. The Sphere-2DPattern has all its reconstructed 3D points within the 3 pixel error. The DLT-like and Distortion-2DPattern approaches show similar results with one 3D reconstruction error within the 5 mm error. The worst result is given by the Sphere-Lines approach with maximum reconstruction error of 8 mm. We observe that the reprojection error for all the methods is below 2 pixels.

5.6. Distribution of the calibration points in the Catadioptric Image

The accuracy of the computed calibration relies on the area occupied by the calibration points in the calibration images. In this experiment we show the importance of the distribution of the calibration points inside the catadioptric images. We define the area to be used by selecting those points closer than r pixels from the image center. The system to calibrate is a central hyper-catadioptric system. Since the images used by each approach are not the same, and also the approaches use different features (points and lines), a full



(a) (b)

Figure 11: Images used in the SfM experiment with reprojected points superimposed (fish-eye).

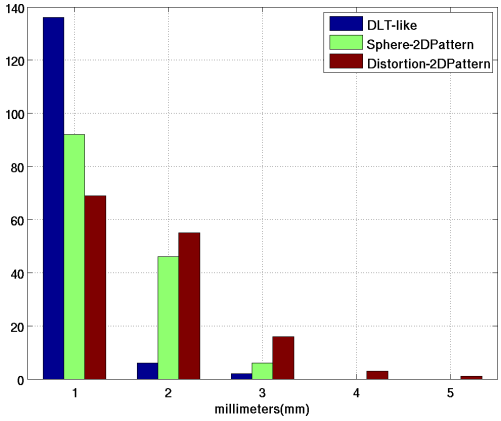


Figure 12: Number of reconstructed 3D points within an error distance in millimeters using a fish-eye lens.

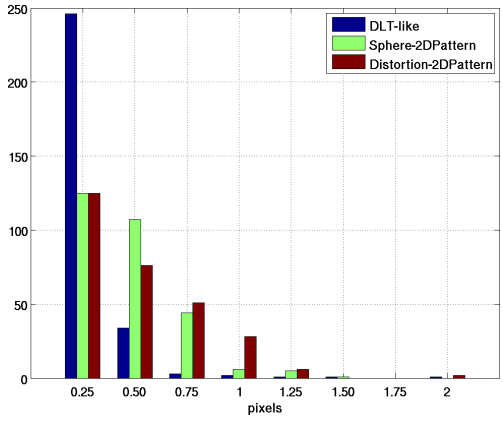


Figure 13: Number of reprojected points within an error distance in pixels using a fish-eye lens.

comparison using the same distances for all the approaches is not possible. Since the approaches Distortion-2DPattern and Sphere-2DPattern share the calibration images set we present their corresponding results together. In the case of DLT-like and Sphere-2DPattern the results are shown separately. The radii r were chosen depending on the requirements of each approach to calibrate the catadioptric system.

5.6.1. Distortion-2DPattern and Sphere-2DPattern

These two approaches require several images to perform the calibration. We select the calibration points that lie closer than r pixels from the image center in all the images of the set in which these points exist. An example for two different radii r can be observed in Fig. 22. In Fig. 23 we show the mean of the reconstruction error for each calibration performed with the points within the areas described by the radii r.

	ξ	(u_0, v_0)
Ground Truth	0.9662	(511.88,399.25)
Sphere-2DPattern	0.8463	(519.14,407.87)
Sphere-Lines	1.00	(537.50,409.82)
DLT-like	0.8819	(525.05,411.89)

Table 4: Comparison of the physical parameters given by the 3 methods based on the sphere model in the non-central system.

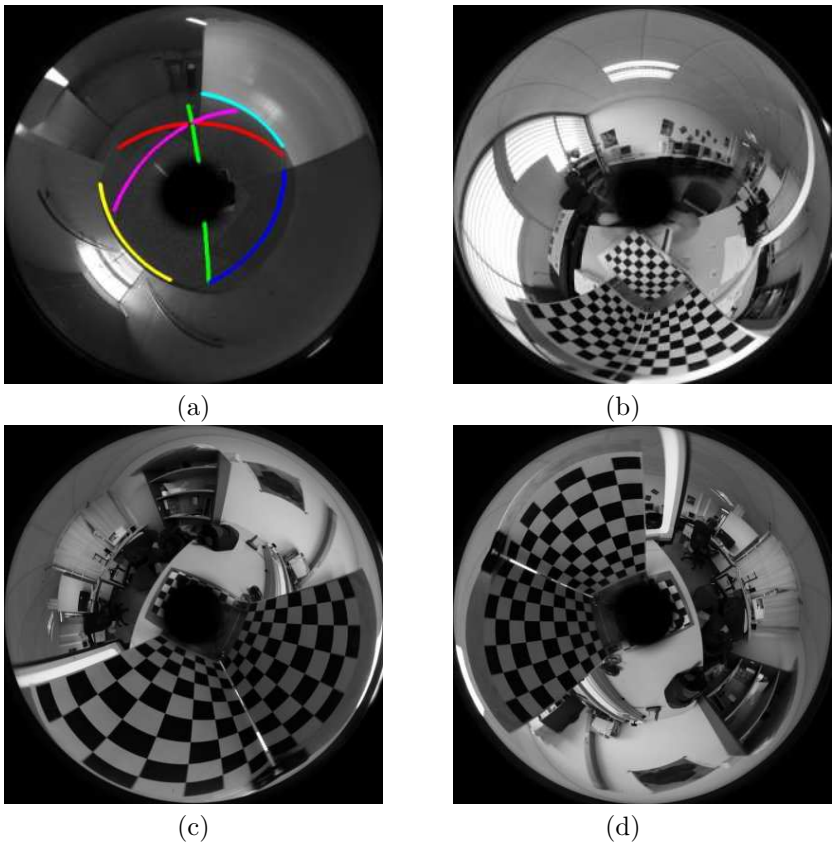


Figure 14: Some images used to calibrate the unknown-shape catadioptric system.(a) Sphere-Lines. (b) DLT-like approach. (c,d) Sphere-2DPattern and Distortion-2DPattern approaches.

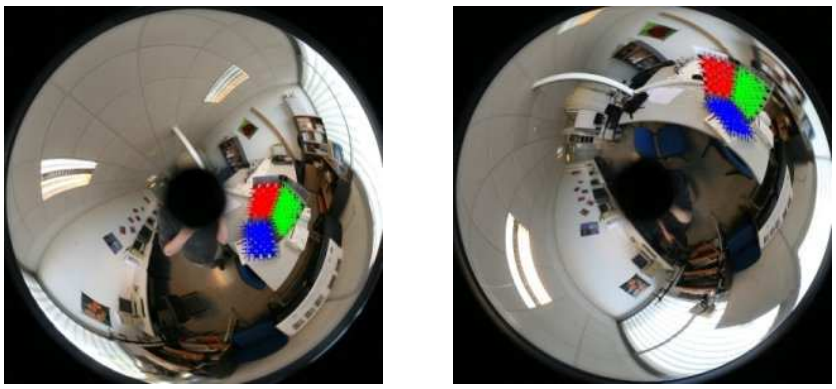


Figure 15: Images used in the SfM experiment with reprojected points superimposed (unknown-shape).

We observe that both approaches give similar results. When the area is small, the number of points decreases and the estimation is worse. The Distortion-2DPattern has an error of 4 mm and the Sphere-2DPattern 2 mm in the worst case. This behavior can be explained by the fact that Sphere-2DPattern estimates the image center from data given by the user and the Distortion-2DPattern does it automatically, depending more on the distribution of the points. In Fig. 24 we show that reprojection error that is under 1 pixel error for both approaches, even with the smallest r.

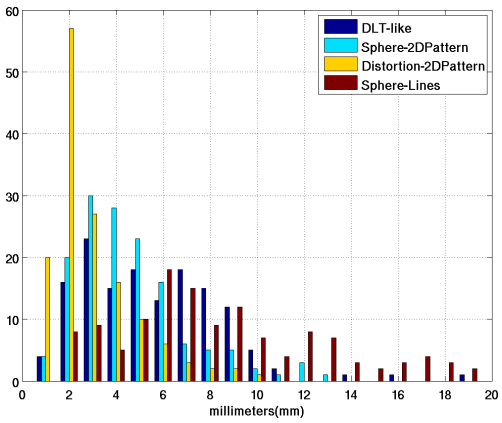


Figure 16: Number of reconstructed 3D points within an error distance in millimeters using an unknown-shape system.

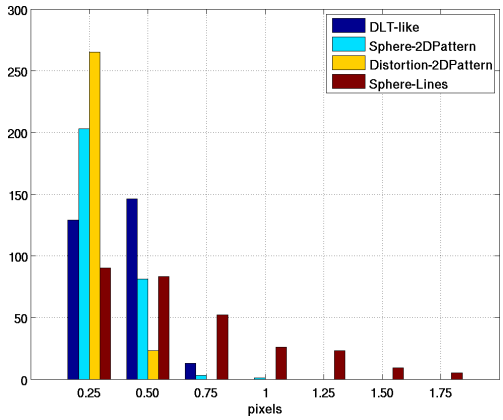


Figure 17: Number of reprojected points within an error distance in pixels using an unknown-shape system.

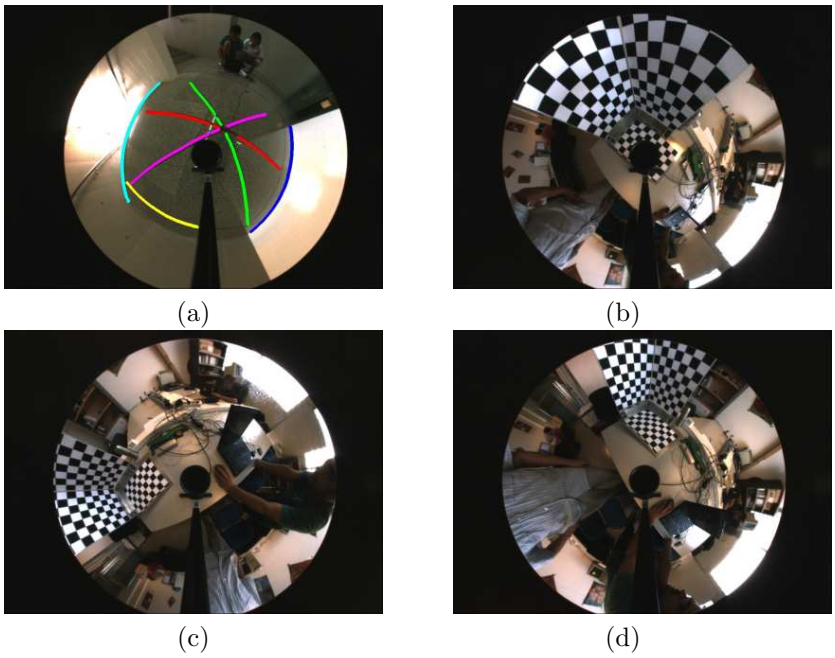


Figure 18: Some images used to calibrate the non-central catadioptric system.(a) Sphere-Lines. (b) DLT-like approach. (c,d) Sphere-2DPattern and Distortion-2DPattern approaches.

5.6.2. DLT-like

Since the DLT-like approach only requires one single image, the points are selected directly using different radii r. In Fig. 25 we show two examples of the points selected for two different radii. In Fig. 26 and Fig. 27 we show the 3D reconstruction error and the reprojection error, respectively. We observe a similar behavior to the previous approaches using large radii. With small radii the results are worse, since with small radii only very few points of the second and third planes are considered (see Fig. 25(a)).

5.6.3. Sphere-Lines

This approach also requires a single image. The image must contain at least 3 line images. The radii limit the length of the lines used to calibrate the system. An example can be observed in Fig. 28. We

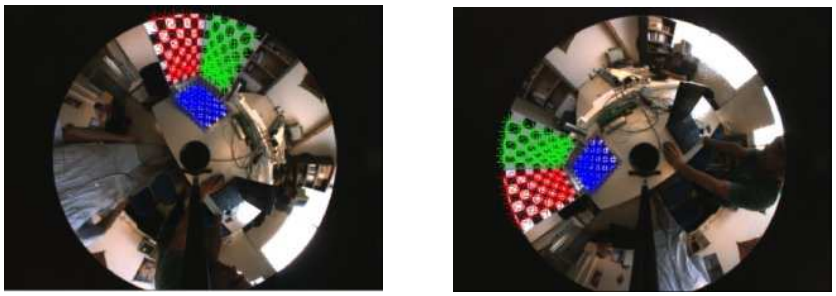


Figure 19: Images used in the SfM experiment with reprojected points superimposed (non-central hyper-catadioptric).

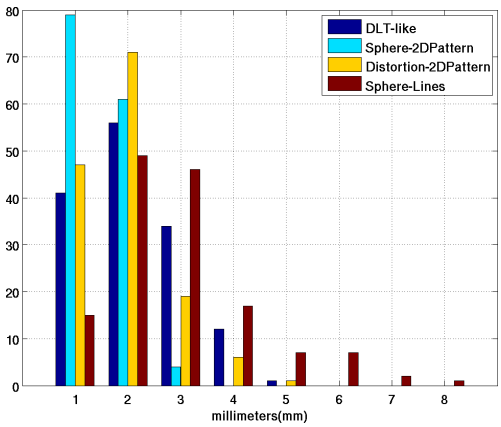


Figure 20: Number of reconstructed 3D points within an error distance in millimeters using a non-central hyper-catadioptric system.

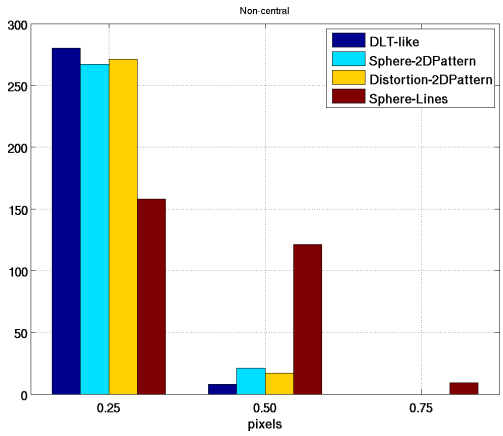


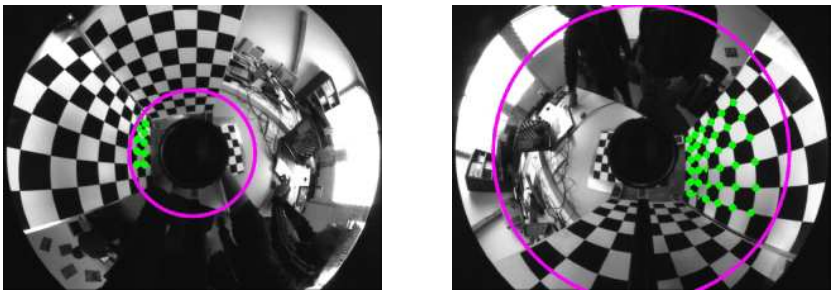
Figure 21: Number of reprojected points within an error distance in pixels using a non-central hyper-catadioptric system.

particularly observed that this approach is quite sensitive to the length and the position of the lines. We show the results where the calibration was possible in the corresponding radii. Fig. 29 and Fig. 30 show the 3D reconstruction error and the reprojection error, respectively. We observe a similar behavior to the other approaches, but having a bigger error, both in the 3D reconstruction error and the reprojection error.

One may think that a comparison of this method using just a few representative elements, in this case lines, present in one single image, against others where hundreds of representative elements (2D points), are extracted from several images, might be unfair. In this order we tried to calibrate the central catadioptric systems using as many lines as possible, present in the same images of the 3D pattern used to calibrate the system using the other approaches. A few examples of the lines used are shown in Fig. 31. The results calibrating the central catadioptric systems using this method did not succeed. We obtained several types of errors and sometimes convergence problems. Because of that we calibrate the central catadioptric systems using the maximum number of lines present in one single image, different from the ones used by the other approaches.

5.7. Discussion

After all these experiments with different systems we observe that all the approaches give good results, with the exception of the Sphere-Lines approach with the fish-eye system, basically because this approach is not designed to deal with such systems. In particular, for the fish-eye lens, the best calibration was achieved with the DLT-like approach. In the case of the Unknown-shape camera, the Distortion-2DPattern approach provided the best result. With respect to the non-central hyper-catadioptric, DLT-like, Sphere-2DPattern and Distortion-2DPattern approaches all gave similar accuracy. Finally, the hyper-catadioptric system was



(a) (b)

Figure 22: Points contained inside the areas described by (a) $r = 150$ pixels and (b) $r = 400$ pixels.

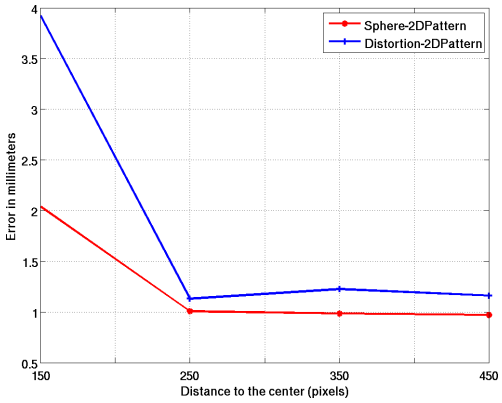


Figure 23: 3D reconstruction error in millimeters. The horizontal axis represents the different radii r.

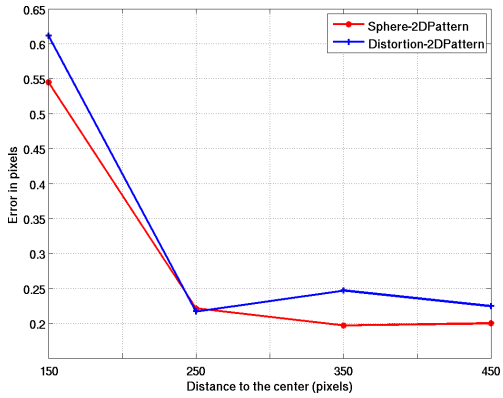


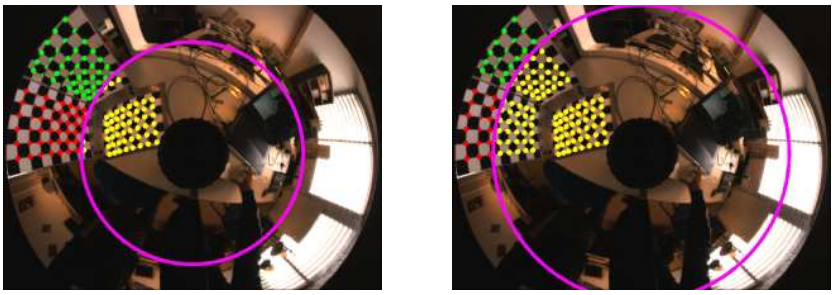
Figure 24: Reprojection error in pixels. The horizontal axis represents the different radii r.

calibrated slightly better by both the DLT-like and the Sphere-2DPattern approaches. We also analyze the importance of the area occupied by the calibration elements (points, lines) in the calibration image(s). All approaches require this area to be as big as possible to compute a good calibration. In the particular case of the Sphere-Lines approach the lines must be as large as possible and must intersect far from the image center. In terms of computing performance, all these approaches perform a non-linear step after the initialization of the intrinsic and extrinsic parameters is computed. The DLT-like approach is the fastest since it estimates less parameters. Next are the Sphere-2DPattern and the Distortion-2DPattern with several extrinsic parameters corresponding to each image to compute, plus the intrinsic parameters. The slowest method is the Sphere-Lines approach, since it uses a complex geometry to compute the self-polar triangle and takes into account every single pixel contained in the line images.

We also consider 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 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. All the image points in the 3D pattern images are given



(a) (b)

Figure 25: Points contained inside the areas described by (a) $r = 300$ pixels and (b) $r = 400$ pixels.

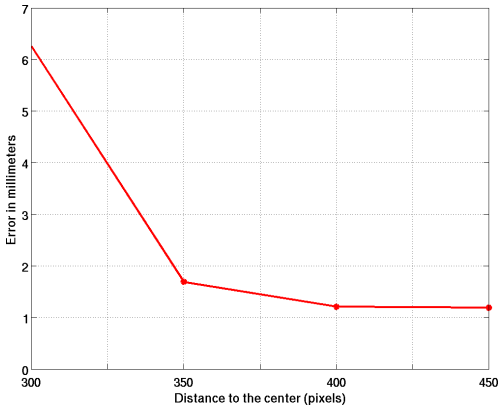


Figure 26: 3D reconstruction error in millimeters. The horizontal axis represents the different radii r.

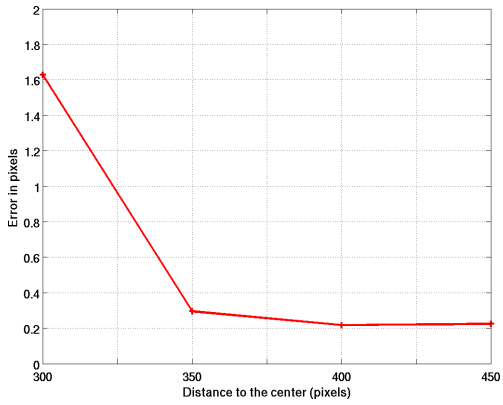


Figure 27: Reprojection error in pixels. The horizontal axis represents the different radii r.

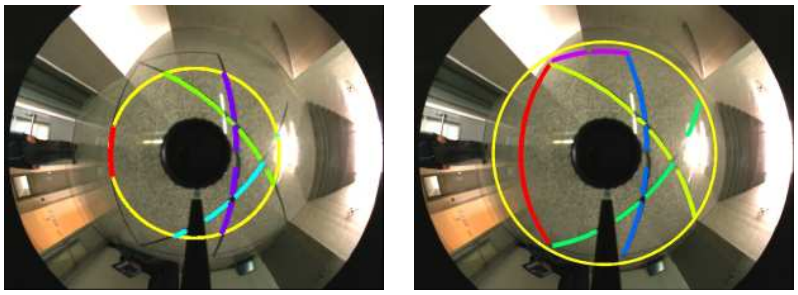
manually. We observed that depending on the image used the mirror parameter ξ is better or worse estimated. Something similar happens with the Sphere-2DPattern approach.

- *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 strongly depends on the principal point estimation. If this estimation is not accurate enough the calibration is not performed properly. 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. Sometimes using more than three lines we had convergence problems or we obtained calibrations containing non-real solutions.

Notice that each method has its own manner to extract the points from the images. In this order we should decouple the matching process from the parameter estimation process.

6. Conclusions

In this paper we have presented a 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 of the omnidirectional 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-eyes. 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



(a) (b)

Figure 28: Points contained inside the areas described by (a) $r = 230$ pixels and (b) $r = 300$ pixels.

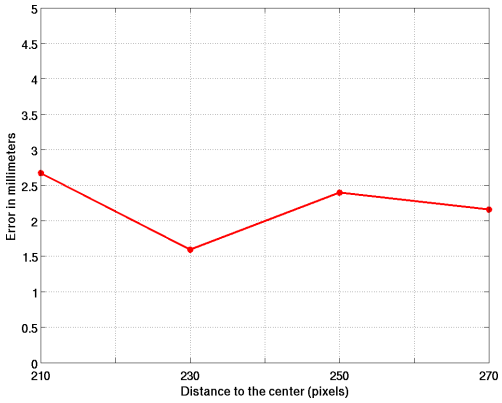


Figure 29: 3D reconstruction error in millimeters. The horizontal axis represents the different radii r.

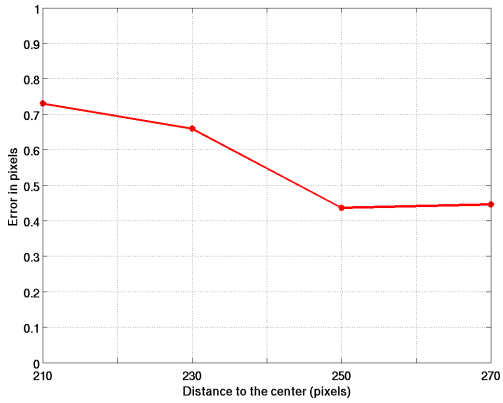


Figure 30: Reprojection error in pixels. The horizontal axis represents the different radii r.

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 useful criteria. This experiment showed that the calibration reached by any of these methods can 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 of the omnidirectional image and mainly in the peripheric area.

7. Acknowledgments

This work was supported by project VISPA DPI2009-14664-C02-01.

References

[1] S. Baker, S. Nayar, A theory of single-viewpoint catadioptric image formation, Int. J. Comput. Vision 35(2) (1999) 175–196.
[2] P. Sturm, S. Ramalingam, J.-P. Tardif, S. Gasparini, J. Barreto, Camera Models and Fundamental Concepts Used in Geometric Computer Vision, Foundations and Trends in Computer Graphics and Vision 6 (1-2) (2011) 1–183.
[3] T. Svoboda, T. Pajdla, Epipolar Geometry for Central Catadioptric Cameras, Int. J. Comput. Vision 49 (1) (2002) 23–37.
[4] D. Strelow, J. Mishler, D. Koes, S. Singh, Precise omnidirectional camera calibration, Proceedings of the Conference on Computer Vision and Pattern Recognition, CVPR 2001. 1 (2001) I–689–I–694 vol.1, ISSN 1063-6919.
[5] C. Geyer, K. Daniilidis, A Unifying Theory for Central Panoramic Systems and Practical Applications, in: Proceedings of the 6th European Conference on Computer Vision-Part II, London, UK, ISBN 3-540-67686-4, 445–461, 2000.
[6] J. P. Barreto, H. Araujo, Issues on the Geometry of Central Catadioptric Image Formation, Proceedings of the Conference on Computer Vision and Pattern Recognition, CVPR 2001. 2 (2001) 422, ISSN 1063-6919.
[7] X. Ying, Z. Hu, Catadioptric camera calibration using geometric invariants, IEEE Transactions on Pattern Analysis and Machine Intelligence 26 (10) (2004) 1260–1271, ISSN 0162-8828.

(a) (b) (c)

Figure 31: Different configurations tested to calibrate the central catadioptric systems using the Sphere-Lines approach. (a) Unknown-shape system, (b) Fish-eye lens and (c) a combination of two images to cover the whole omnidirectional image with lines.

[8] R. Swaminathan, S. Nayar, Non-metric calibration of wide-angle lenses and polycameras, in: Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on., vol. 22, 1172–1178, 1999.

[9] B. Micusik, T. Pajdla, Estimation of omnidirectional camera model from epipolar geometry, in: Computer Vision and Pattern Recognition., vol. 1, ISSN 1063-6919, I–485 – I–490 vol.1, 2003.

[10] J. Kannala, S. Brandt, A generic camera calibration method for fish-eye lenses, in: Proceedings of the 17th International Conference on Pattern Recognition. ICPR 2004, vol. 1, ISSN 1051-4651, 10–13 Vol.1, 2004.

[11] J. Courbon, Y. Mezouar, L. Eck, P. Martinet, A generic fisheye camera model for robotic applications, in: Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on, 1683–1688, 2007.

[12] M. Grossberg, S. Nayar, A general imaging model and a method for finding its parameters, in: Proceedings of the Eighth IEEE International Conference on Computer Vision, ICCV 2001., vol. 2, 108–115 vol.2, 2001.

[13] S. Ramalingam, P. Sturm, S. Lodha, Towards complete generic camera calibration, in: Proceedings of the Conference on Computer Vision and Pattern Recognition. CVPR 2005., vol. 1, ISSN 1063-6919, 1093–1098 vol. 1, 2005.

[14] O. Morel, D. Fofi, Calibration of Catadioptric Sensors by Polarization Imaging, IEEE International Conference on Robotics and Automation, (2007) 3939–3944ISSN 1050-4729.

[15] D. Scaramuzza, A. Martinelli, R. Siegwart, A Flexible Technique for Accurate Omnidirectional Camera Calibration and Structure from Motion, Proceedings of IEEE International Conference on Computer Vision Systems, 2006.

[16] C. Mei, P. Rives, Single View Point Omnidirectional Camera Calibration from Planar Grids, in: IEEE International Conference on Robotics and Automation, ISSN 1050-4729, 3945–3950, 2007.

[17] X.-M. Deng, F.-C. Wu, Y.-H. Wu, An Easy Calibration Method for Central Catadioptric Cameras, Acta Automatica Sinica 33 (8) (2007) 801 – 808, ISSN 1874-1029.

[18] O. Frank, R. Katz, C. Tisse, H. Durrant Whyte, Camera calibration for miniature, low-cost, wide-angle imaging systems, in: British Machine Vision Conference, 2007.

[19] S. Gasparini, P. Sturm, J. P. Barreto, Plane-based calibration of central catadioptric cameras, in: Computer Vision, 2009 IEEE 12th International Conference on, ISSN 1550-5499, 1195–1202, 2009.

[20] L. Puig, Y. Bastanlar, P. Sturm, J. J. Guerrero, J. Barreto, Calibration of Central Catadioptric Cameras Using a DLT-Like Approach, International Journal of Computer Vision 93 (2011) 101–114, ISSN 0920-5691.

[21] Y. Bastanlar, L. Puig, P. Sturm, J. J. Guerrero, J. Barreto, DLT-Like Calibration of Central Catadioptric Cameras, in: Workshop on Omnidirectional Vision, Camera Networks and Non-Classical Cameras, Marseille, France, 2008.

[22] C. Toepfer, T. Ehlgen, A Unifying Omnidirectional Camera Model and its Applications, in: Proceedings of the 11th International Conference on Computer Vision, ICCV 2007., ISSN 1550-5499, 1–5, 2007.

[23] C. Geyer, K. Daniilidis, Paracatadioptric camera calibration, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (5) (2002) 687–695, ISSN 0162-8828.

[24] C. Geyer, K. Daniilidis, Catadioptric camera calibration, in: Proceedings of the 7th IEEE International Conference on Computer Vision, vol. 1, 398–404 vol.1, 1999.

[25] J. P. Barreto, H. Araujo, Geometric Properties of Central Catadioptric Line Images, in: Proceedings of the 7th European Conference on Computer Vision-Part IV, ECCV '02, London, UK, ISBN 3-540-43748-7, 237–251, 2002.

[26] J. P. Barreto, H. Araujo, Geometric Properties of Central Catadioptric Line Images and Their Application in Calibration, IEEE Transactions on Pattern Analysis and Machine Intelligence 27 (8) (2005) 1327–1333.

[27] X. Ying, H. Zha, Simultaneously calibrating catadioptric camera and detecting line features using Hough transform, IEEE/RSJ International Conference on Intelligent Robots and Systems, 2005, IROS 2005. (2005) 412–417.

[28] B. Vandeportaele, M. Cattoen, P. Marthon, P. Gurdjos, A New Linear Calibration Method for Paracatadioptric Cameras, Proceedings of the 18th International Conference on Pattern Recognition. ICPR 2006. 4 (2006) 647–651.

[29] Y. Wu, Y. Li, Z. Hu, Easy Calibration for Para-catadioptric-like Camera, IEEE/RSJ International Conference on Intelligent Robots and Systems (2006) 5719–5724.

[30] V. Caglioti, P. Taddei, G. Boracchi, S. Gasparini, A. Giusti, Single-Image Calibration of Off-Axis Catadioptric Cameras Using Lines, Proceedings of the 11th IEEE International Conference on Computer Vision, ICCV 2007. (2007) 1–6.

[31] F. Wu, F. Duan, Z. Hu, Y. Wu, A new linear algorithm for calibrating central catadioptric cameras, Pattern Recognition

41 (10) (2008) 3166–3172, ISSN 0031-3203.
[32] S. B. Kang, Catadioptric self-calibration, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, 201–207, 2000.
[33] B. Micusik, T. Pajdla, Structure from Motion with Wide Circular Field of View Cameras, IEEE Trans. Pattern Anal. Mach. Intell. 28 (7) (2006) 1135–1149, ISSN 0162-8828.
[34] S. Ramalingam, P. Sturm, S. K. Lodha, Generic self-calibration of central cameras, Computer Vision and Image Understanding 114 (2) (2010) 210 – 219, ISSN 1077-3142, special issue on Omnidirectional Vision, Camera Networks and Non-conventional Cameras.
[35] F. Espuny, J. Burgos Gil, Generic Self-calibration of Central Cameras from Two Rotational Flows, International Journal of Computer Vision 91 (2011) 131–145, ISSN 0920-5691, 10.1007/s11263-010-0335-9.
[36] R. Tsai, A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses, Robotics and Automation, IEEE Journal of 3 (4) (1987) 323 –344, ISSN 0882-4967.
[37] P. Vasseur, E. M. Mouaddib, Central Catadioptric Line Detection, in: British Machine Vision Conference, 2004.
[38] D. Aliaga, Accurate catadioptric calibration for real-time pose estimation in room-size environments, Proceedings of the 8th IEEE International Conference on Computer Vision, ICCV 2001. 1 (2001) 127–134 vol.1.
[39] Y. Wu, Z. Hu, Geometric Invariants and Applications under Catadioptric Camera Model, IEEE International Conference on Computer Vision 2 (2005) 1547–1554, ISSN 1550-5499.
[40] P. Sturm, S. Ramalingam, A Generic Concept for Camera Calibration, in: T. Pajdla, J. Matas (Eds.), Computer Vision - ECCV 2004, vol. 3022 of *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, 1–13, 2004.
[41] J. P. Barreto, K. Daniilidis, Epipolar Geometry of Central Projection Systems Using Veronese Maps, in: Computer Vision and Pattern Recognition, IEEE Computer Society, Washington, DC, USA, ISBN 0-7695-2597-0, 1258–1265, 2006.
[42] P. Sturm, Mixing Catadioptric and Perspective Cameras, in: Workshop on Omnidirectional Vision, Copenhagen, Denmark, 37–44, 2002.
[43] J. Heikkila, O. Silven, A four-step camera calibration procedure with implicit image correction, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1106–1112, 1997.
[44] R. I. Hartley, A. Zisserman, Multiple View Geometry in Computer Vision, Cambridge University Press, ISBN: 0521623049, 2000.