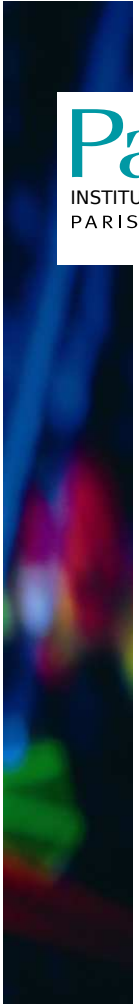
Direct self-calibration of central catadioptric omnidirectional cameras

Adan Salazar-Garibay



ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY

MINES ParisTech

École doctorale n°84 :
Sciences et technologies de l'information et de la communication

Doctorat ParisTech

T H È S E

préparée à l'INRIA Sophia Antipolis

pour obtenir le grade de docteur délivré par

l'École Nationale Supérieure des Mines de Paris

Spécialité « Informatique temps-réel, robotique et automatique »

présentée et soutenue publiquement par

Adan SALAZAR-GARIBAY

en septembre 2011

sur le thème en anglais

Direct Self-Calibration of Central Catadioptric Omnidirectional Cameras

Directeur de thèse : **Ezio MALIS**

Jury

El Mustapha MOUADDIB, Professeur, Laboratoire MIS, UPJV	Rapporteur
Peter STURM, Directeur de Recherche, PERCEPTION, INRIA Grenoble – Rhône-Alpes	Rapporteur
Joao BARRETO, Professeur, Université de Coimbra, Portugal	Examinateur
Patrick RIVES, Directeur de recherche, AROBAS, INRIA Sophia Antipolis-Méditerranée	Examinateur
Jean-Paul MARMORAT, Directeur de recherche, École des Mines de Paris	Examinateur
Ezio MALIS, Directeur de recherche, EVOLUTION, INRIA Sophia Antipolis-Méditerranée	Examinateur

T H È S E

MINES ParisTech
Centre de Mathématiques Appliquées
Rue Claude Daunesse B.P. 207, 06904 Sophia Antipolis Cedex, France

Abstract

Calibrating a camera means determining the geometric properties of the imaging process, i.e., the transformation that maps a three-dimensional point, expressed with respect to a reference frame, onto its two-dimensional image whose coordinates are expressed in pixel units. This process is required when recovering 3D information. More precisely, we have to know the translation and rotation of the visual sensor with respect to the rest of the frame system (extrinsic parameters), and the different parameters of the lens, such as focal length, magnitude factors and the optical center retinal location (intrinsic parameters).

Although the camera calibration problem is well understood, no method allowing the robust direct on-line self-calibration for any central omnidirectional camera is known. Existing self-calibration techniques attempt to calibrate from point correspondences, lines, circles or a specific camera motion. Even though interesting results can be achieved, self-calibration still suffers from some limitations such as small number of feature points, difficult detection of lines, undesirable camera motion and taking into account a specific mirror. Consequently, the aim of this thesis is to propose a new algorithm that overcomes these limitations and that can be adopted by any robotic application or by any other practical implementation in which the calibration process is not straightforward. This algorithm works directly with the image intensity, makes the minimum of assumptions about the particular structure of the scene being viewed, stays valid for all central catadioptric systems and needs no prior knowledge about extrinsic and intrinsic parameters.

Also, part of this thesis is dedicated to formalizing the uniqueness of the solution for the calibration problem of central catadioptric omnidirectional cameras. For the greatest part of the work on omnidirectional camera calibration it has been observed that, in the case of a non-planar mirror, two images acquired from different points of view suffice to calibrate an omnidirectional camera. However, to our knowledge, no theoretical proof of the uniqueness of the solution has been provided yet. In this thesis the calibration problem is formalized by using a unified model that is valid for all central catadioptric omnidirectional cameras. It is also shown that the uniqueness of the problem can be derived from the solution of non-linear equations. However, due to the complexity of the non-linear equations to be solved for the general case, this thesis devises the uniqueness of the solution for the particular case of a parabolic mirror when the camera's movement is a rotation plus a translation.

To my late father and to my late brothers

To my mother and brother

Contents

Chapter 1

Introduction

Engineers, scientists, and business men are increasingly turning towards nature for design inspirations. The reason is simple. Nature, through billions of years of trial and error, has produced effective solutions to innumerable complex real-world problems. The rigorous competition of natural selection means waste and efficiency are not tolerated in natural systems, unlike many of the technologies devised by humans. When scientits wished to build a device to obtain a large field of view, they started by asking the question: Are there any existing systems with such capabilities? The answer is yes. In fact, nature offers a wide range of examples - in particular in the insect kingdom (ants, spiders, houseflies, etc.). For instance, the vision system of insects equipped with compound eyes are considered as omnidirectional sensors (See Figure 1.1). Compound eyes are found among the arthropods and are composed of many simple facets, which depending on the details of anatomy, may give either a single pixelated image or multiple images, per eye. Each sensor has its own lens and photosensitive cell. Some eyes have up to 28,000 such sensors, which are arranged hexagonally, and which can give a full 360 degrees field of view.

Figure 1.1: *Housefly's compound eye.*

1.1 Omnidirectional Vision in Science

In science, the term omnidirectional vision refers to vision sensors with a very large field of view (sensor with a horizontal field of view of 360 degrees and a variable vertical field of view usually between 60 and 150 degrees). Omnidirectional vision involves capture and automatic interpretation of images usually depicting full 360 degrees (horizontal panoramas) view of the surroundings. Sometimes the field of view band can be aligned vertically, or in any other plane, usually by means of deploying a

circularly symmetrical mirror aligned in the required direction. Hemispherical fields of view (360 degrees by 180 degrees) are also used, usually by deploying a fish-eye lens. Finally, separate views can be combined, thus forming the ideal full spherical projection.

1.1.1 Different sensors for capturing wide fields of view

The modern omnidirectional sensors used in computer vision and robot vision can be devided into three main groups:

1. Panoramic cameras;
2. Compound-eye cameras;
3. Omnidirectional cameras.

1.1.1.1 Panoramic cameras

Panoramic cameras are sensors with the ability to produce perspective panoramic views by using only single camera. Most of these sensors pan a camera around the vertical axis passing through the focal point of the camera (See Figure 1.2). These panoramic cameras provides high definition panoramic images. However, very fine calibration and synchronization of the camera movements are required. Some of the sensors do not have such requirement. Instead of rotating around its focal point, they rotate around a vertical axis at a certain distance from its focal point. Rough range panoramic images are obtained by matching the views of object from different positions. The main drawback is that this technique is very slow and not suitable for use in a dynamic environment. In a dynamic environment, we need cameras capable of capturing a global view in one shot in order to have, at any time in the field of view, all the moving objects (e.g. omnidirectional cameras).

Figure 1.2: *Panoramic camera.*

1.1.1.2 Compound-eye cameras

Compound-eye cameras use multiple cameras to grab pictures from different directions and then stitch them together to produce a global view of the environment (Figure 1.3). The advantages of these cameras are the high resolution that can be achieved and the possibility of grabbing pictures from different directions at the same time. However, the disadvantage is the complexity of the system and the need of an accurate calibration process.

Figure 1.3: *Compound-eye camera.*

1.1.1.3 Omnidirectional cameras

Omnidirectional cameras are sensors that capture light from all directions covering a full sphere. In practice, however, most omnidirectional cameras cover only approximately a semi-sphere, or the full 360 degrees along the equator of the sphere but exclude the poles of the sphere.

Omnidirectional cameras can be central or non-central. A camera is central, or has a single viewpoint, if all rays intersect in a single point. The rays may be completely independent for an arbitrary camera, in particular, they do not have to intersect in a single viewpoint. The reason why a single viewpoint is so desirable is because the projection models, and therefore the algorithms, are simple. Also, this property allows to generate pure perspective images from the observed ones.

Omnidirectional cameras that use only the refractive effect of the lens to inflect the light are called dioptric cameras[1]. Cameras that use the combined effects of reflection from a mirror and refraction from a lens are called catadioptric cameras.

[1]Dioptrics is the study of the refraction of light by lenses. Telescopes that create their image with an objective that is a convex lens (refractors) are said to be "dioptric" telescopes.

Omnidirectional cameras can be grouped into two categories: cameras that use a special lens and cameras that use a convex mirror and a set of lenses.

Special lens cameras. Special lens cameras (also called fish-eye cameras) are imaging systems combining a fish-eye lens and a conventional camera[2]. Thanks to the fish-eye lens, they can acquire almost a hemispherical view (See Figure 1.4). Their main advantage with respect to catadioptric sensors is that they do not exhibit a dead area. The drawback is that the resolution of the images is good at the centre but very low at the periphery. This is not good for robot navigation, where the objects to locate lie on the floor and they appear at the horizon or below. In other words, the resolution is very good at the ceiling but poor at the horizon.

Figure 1.4: *Special lens camera. Left. Fish-eye camera. Right. Fish-eye shot.*

Convex mirror cameras. Convex mirror cameras or catadioptric cameras are the most widely used in robotics to obtain omnidirectional images. The sensor is composed by a perspective camera pointed upward to the vertex of a convex mirror (See Figure 1.5). The optical axis of the camera and the geometrical axis of the mirror are aligned. This system is usually fixed on top of a mobile robot. Mirrors with different shapes can be used. The most common are: conical, hemispherical, hyperboloidal and paraboloidal. Every shape presents different properties that one has to take into account when choosing the mirror for a particular task.

Sometimes catadioptric cameras tend to be big in comparison to conventional cameras. This is due to the fact that capturing a wide unobstructed field of view requires the lens and the mirror to be adequately separated from each other. To overcome this limitation, folded catadioptric cameras (cameras combining several mirrors and

[2]In photography, a fish-eye lens is a wide-angle lens that takes in an extremely wide, hemispherical image. Originally developed for use in meteorology to study cloud formation and called *whole-sky lenses*, fish-eye lenses quickly became popular in general photography for their unique, distorted appearance. They are often used by photographers shooting broad landscapes to capture the curvature of the Earth.

a set of lens) were developed. They use the optical folding method to fold the optical path between the curved mirror and the lens system. Folding with a curved mirror creates an 180 degrees fold and can reduce undesired optical effects.

Figure 1.5: **Convex mirror camera.** *Left.* A convex mirror camera or catadioptric camera composed by a perspective camera and a parabolic mirror. *Right.* Omnidirectional images taken with a catadioptric camera.

1.1.2 Omnidirectional vision in computer vision and robotics

Omnidirectional vision is a technique for capturing the world in an image. It is not mature yet for commercial applications, but its use is growing quickly. More and more laboratories around the world are developing new sensors and new applications for or with omnidirectional vision. Applications include all situations where large fields of view are needed, such as video-conferencing, site modeling, wide area surveillance, virtual reality, navigation and guidance of (autonomous) vehicles, tracking and motion detection, visual odometry and simultaneous localization and mapping (SLAM), among others [DGSW00, RCS02, Gre85, BS, Yag99, BK01, EMW01].

Due to its ability to capture a 360 degree view, omnidirectional vision has found the largest application in the domain of mobile robotics (Figure 1.6 illustrates a mobile robot with a catadioptric panoramic camera). It is easier to recognize previously observed places. Also, it is much easier to deal with the rotation of the camera mounted on a robot, as this will not make objects disappear from view but only change their positions in the image. Egomotion estimation algorithms are simpler when using omnidirectional images, since the rotation and translation components of a movement can be decoupled. Also, the accuracy of self-localisation is largely influenced by the locations of the landmarks relative to the robot. Only omnidirectional vision allows selection of landmarks all around the robot. Additionally, omnidirectional vision lends itself to a particularly simple stereo epipolar geometry and consequently to instantaneous range finding that is not limited by a narrow field of view. As we can see, omnidirectional cameras offer a number of significant benefits. Nevertheless, their practical use is often burdened by the calibration phase.

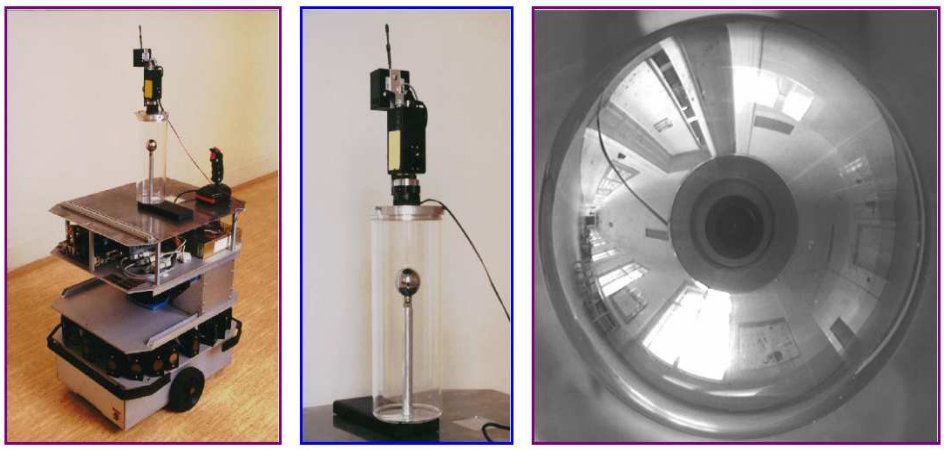


Figure 1.6: **Mobile robot** *Left.* The mobile robot with a catadioptric panoramic camera. *Middle.* The detail view of the camera. *Right.* The panoramic image shows how the camera sees itself and occludes a part of the ceiling.

1.2 Camera calibration

Camera calibration is a way of examining several images, or a video, and deducing what the camera situation was at the time the images were captured. Camera calibration is used primarily in robotic applications, and when modeling scenes virtually based on real input.

One of the main uses of camera calibration is to determine where a camera was in relation to a scene in an image. Camera calibration achieves this, by using formulas to essentially work backwards, and deduce where the camera was relative to the scene.

Camera calibration can also be used to figure out other things about the camera in relation to the scene. For example, we can infer the focal length used to photograph the scene. We can also deduce the skew factor of the image, and any lens distortion that may have been introduced, creating a barrel effect, pincushion effect or mustache effect. We can also deduce whether the actual camera pixels were square, and the horizontal and vertical scaling factors for the pixels might have been.

We can also use camera calibration to take an image sent to a computer, and deduce where various coordinates are in the real world. This type of deduction is crucial to the functioning of robots that are meant to interact visually with the physical world. These robots can then use an image, or video input device and calibrate in order to figure out where objects it sees might actually be in the real world, in actual terms of distance and vector.

This is one of the major areas of study in robotics, as faster, more accurate, methods of calibrating allow robots to interact with the world in more sophisticated ways. A robot with a poor ability to discern the distance of objects will have to rely largely on trial and error to move over terrain or manipulate an object, whereas one that is able to accurately model its own place in the world in relation to other objects, is able to move seamlessly in the world.

Many approaches to estimate the calibration of catadioptric omnidirectional cameras have been proposed. Those approaches can be divided into two main groups. The *first group* includes methods which exploit prior knowledge about the scene, such as the presence of calibration patterns [GSB09, SB08, BPS+08, MBMR06, SMS06, DWW07, CBDD00], lines [GD02, BA05, YZ05, WLH06], points [Vas04, Ali01, WH05], sphere images [YH03, XY05] or polarized images [MSF07]. The *second group* covers techniques that do not use this knowledge. This includes calibration methods from pure rotation [BP02] or planar motion of the camera [GN98] and self-calibration procedures [GD01, Kan00, MP04, FTB02, SR04, BV08].

1.3 Motivation

Although the camera calibration problem is well understood, no method allowing the robust on-line self-calibration for any central omnidirectional camera is known. Existing self-calibration techniques attempt to calibrate from point correspondences, lines, circles or a specific camera motion. Even though interesting results can be achieved, self-calibration still suffers from some limitations such as small number of features points, difficult detection of lines, undesirable camera motion and taking into account of a specific type of mirror. Consequently, the motivation of this thesis is to propose a new algorithm that overcomes these limitations. This algorithm works directly with the image intensity, makes the minimum of assumptions about the particular structure of the scene being viewed, stays valid for all central catadioptric systems and needs no prior knowledge about extrinsic and intrinsic parameters.

1.4 Goal of the thesis

The aim of this thesis is to develop an on-line self-calibration algorithm for the simplification of the calibration process required by any central omnidirectional camera. This method can be adopted by any robotic application or by any other practical implementation in which the calibration process is not straightforward.

1.5 Methodology

In order to achieve the main goal of this thesis, the following methodology is applied: Firstly, a unified model that is valid for all central omnidirectional cameras is

adopted. This stage is followed by an algorithm that accomplishes the visual tracking of a plane. Experiments with synthetic as well as real data are performed to validate the visual tracking method. Secondly, the visual tracking algorithm is applied to the on-line self-calibration of the catadioptric omnidirectional sensor. Finally, different catadioptric omnidirectional sensors are employed to test the proposed algorithm.

1.6 Contributions of this thesis

The main contribution of this thesis is to show that an on-line self-calibration method based only on planar tracking works for any central omnidirectional camera in a robust manner. This can be very helpful in any application where camera calibration is impossible or hard to obtain. Further, the following are proposed in the thesis:

An algorithm for visual tracking of a plane with an uncalibrated catadioptric camera. An algorithm to efficiently track a plane in an omnidirectional image without requiring the prior calibration of the sensor is proposed. The approach deals with a nonlinear optimization problem that can be solved for small displacements between two images acquired at a video rate.

Proof of the uniqueness of the solution for the calibration of central catadioptric omnidirectional cameras For the greatest part of the work on omnidirectional camera calibration, it has been observed that, in the case of a non-planar mirror, two images acquired from different points of view suffice to calibrate an omnidirectional camera. However, to our knowledge, no theoretical proof of the uniqueness of the solution has been provided yet. Hence, part of this work is dedicated to formalizing the uniqueness of the solution for the calibration problem of central catadioptric omnidirectional cameras.

1.7 Structure of the thesis

Chapter 2 deals with the theory of omnidirectional camera geometry. It describes the omnidirectional camera projection model, the image formation and the camera calibration methods which exploit prior knowledge about the scene. Chapter 3 begins with an overview of tracking approaches in the literature using omnidirectional cameras. Then, an algorithm for efficiently tracking a plane in an omnidirectional image without requiring the prior calibration of the sensor is presented. Chapter 4 focuses on the main goal of this thesis, the on-line self-calibration of central catadioptric omnidirectional cameras. First, a brief state of the art of self-calibration methods is presented. Then, the on-line self-calibration approach is shown and tested with synthetic data and real data for different types of omndirectional cameras. Finally, the mathematical proof to the uniqueness of the solution for the calibration problem of a central para-catadioptric omnidirectional camera is presented. General conclusions and directions for future work are given in Chapter 5. Theorical demonstrations are included in the appendices.

Chapter 2

Central Catadioptric Omnidirectional Cameras

This chapter is dedicated to central catadioptric cameras. First, we present a brief overview of catadioptric camera through its history. Then, we describe how to link the image in the sensor to the light rays emitted by a region of space. In this part, we show that wide-angle sensors impose a different view of the world through spherical perspective projection. We also present the unified projection model which uses spherical perspective projection. We focus on this model because it presents a compromise between genericity and over-parameterisation and its parameters are easily identifiable; therefore, we use it throughout this work in order to accomplish all the experiments and to provide reliable results. Finally, we analyse a state of the art of existing calibration methods which exploits prior knowledge about the scene, such as the presence of calibration patterns, lines, points or sphere images.

2.1 Overview

The concept of catadioptric camera already appeared in the presentation of René Descartes in 1637 in *Discours de la Methode* [DS37]. He showed that refractive as well as reflective ovals (conical lenses and mirrors) focus light into a single point. The idea was later re-phrased by Feynman *et al.* in 1963 [RPFS63]. The first catadioptric cameras were composed of lenses rotating around a given axis (swing lens cameras). Because the camera was stationary, the acquired field of view was limited between 120 degrees and 150 degrees. Rotating cameras, created shortly after, did not have this limitation and made it possible to create 360 degrees views of the environment. However, if we can process the images, as is done regularly with computers nowadays, a way of capturing large field of views without any moving parts can be achieved. Two main approaches have been used: (1) a mirror system constructed from several planar mirrors, where a camera is assigned to each mirror plane; (2) one camera observes convex or concave mirror. Rees [Ree70] was the first to patent the combination of a perspective camera and a convex mirror (in this case a hyperbolic mirror). In his US patent, he described how to capture omnidirectional images that can be transformed to correct perspective views (no parallax). Jeffrey R. Charles designed a mirror system for his Single Lens Reflex (SLR) camera. He proposed a darkroom process to transform a panoramic image to a cylindrical projection [Cha76]. It is much later, in the 90's, that omnidirectional vision became an active research topic in computer and robot vision. Chahl and Srinivasan designed a convex mirror with regard to quality imaging [CS97]. Geb proposed an imaging sensor with a spherical mirror for teleoperational navigation on a mobile

vehicle [Geb98]. Kawanishi *et al.* proposed an omnidirectional sensor covering a whole viewing sphere [TKY98]. They designed an hexagonal mirror assembled with 6 cameras. Two such catadioptric cameras were symmetrically connected back to back to observe the whole surroundings. Yagi *et al.* used a conic-shaped mirror. Vertical edges in the panoramic image were extracted and with acoustic sensor cooperation a trajectory of a mobile robot was found [YYY95]. Yamazawa *et al.* detected obstacles using a catadioptric sensor with a hyperbolic mirror [YY00]. Under the assumption of planar motion they computed the distance between the sensor and the obstacle. Their method was based on a scale invariant image transformation. In [DSR96], the authors proposed a catadioptric system with a double lobed mirror. The shape was not precisely defined. Hamit [Ham97] described various approaches to get omnidirectional images using different types of mirrors. Nayar *et al.* presented several prototypes of catadioptric cameras using a parabolic mirror in combination with ortographic cameras [Nay97]. Hicks and Bajcsy presented a family of reflective surfaces that provided a wide field of view while preserving the geometry of a plane perpendicular to their axis of symmetry [HB99]. Their mirror design had the ability to give a normal camera a *birds eye* view of its surroundigs. In the last decade many other systems have been and continue to be designed, as new applications, technological opportunities or research results. For instance, Layerle Jean-Francois [LSEM08] has proposed the design of a new catadioptric sensor using two different mirror shapes for a simultaneous tracking of the driver's face and the road scene. They showed how the mirror design allows the perception of relevant information in the vehicle: a panoramic view of the environment inside and outside, and a sufficient resolution of the driver's face for a gaze tracking.

2.2 Image Formation

By image formation we will understand the formation of a digital image from a surrounding scene through an optical (including mirrors) and a digitization process. This is the basic step for using vision sensors but it is not a straightforward task: sensors are not perfect and we need to model small errors of design.

Let I be an image of the world obtained through an optical device. We will consider I to be a two-dimensional finite array containing intensity values (irradiance). I can be seen as a function:

$$\begin{array}{rccc} I: & \Omega \subset \mathbb{R}^2 & \longrightarrow & \mathbb{R}_+ \\ & (u,v) & \longmapsto & I(u,v) \end{array} \tag{2.1}$$

The irradiance at an image point $\mathbf{p} = (u, v)$ is due to the energy emitted from a region of space determined by the optical properties of the device. In the case of a *central* device with a unique viewpoint, the direction of the energy source is represented by a projective ray (a half-line) with initial point the optical center (or the focus) of the device (noted C in Figures 2.1 and 2.2).

2.3 Projection Models

As we said in the previous chapter, catadioptric omnidirectional cameras are sensors composed of a perspective camera pointed upwards to the vertex of a convex mirror. Therefore, the image formation of this kind of sensor is described by the well known perspective camera model complemented by a non linear part given by a function due to the mirror standing between the world scene and the camera. Let us present how the planar perspective projection model and the spherical perspective projection can be related to model the image formation for this kind of a sensor.

2.3.1 Planar perspective projection

The most common way to model a single camera is to use the well known perspective projection model [Fau93, HZ00]. The standard perspective camera model maps all scene points $\mathbf{X}$ (3D point) from a line passing through an optical center of a camera to one image point $\mathbf{p}$. Figure 2.1 depicts the standard perspective projection model. Let the optical center of the device be the camera reference frame, so point $\mathbf{X}$ becomes $\mathbf{X}_C = (X, Y, Z)^\top$ and its projection is made as follows :

1. The point $\mathbf{X}_C = (X, Y, Z)^\top$ is projected to the normalised plane $\pi_\mathbf{m}$ by the following equation:

$$\mathbf{m} = (x, y, 1)^\top = \begin{pmatrix} \frac{1}{Z} & 0 & 0 \\ 0 & \frac{1}{Z} & 0 \\ 0 & 0 & \frac{1}{Z} \end{pmatrix} (X, Y, Z)^\top = \left(\frac{X}{Z}, \frac{Y}{Z}, 1 \right)^\top$$

2. Let $\alpha_1 = k_u f$ be the horizontal focal length in pixels and $\alpha_2 = k_v f$ the vertical focal length in pixels, where f is the focal length, k_u and k_v are the scaling factors for row pixels and column pixels (camera pixels are not necessarily square). s the skew factor (represents the non-ortogonality between rows and columns of the sensor cells) and (u_0, v_0) the principal point (it is typically not at (width/2, height/2) of image), then the projection of $\mathbf{m}$ in homogeneous coordinates to the image plane $\pi_\mathbf{p}$ is obtained linearly by :

$$\mathbf{p} = (u, v, 1)^\top = \mathbf{P}\mathbf{X}_C = \mathbf{K}\mathbf{m} = \begin{pmatrix} \alpha_1 & s & u_0 \\ 0 & \alpha_2 & v_0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{m}$$

This model is not adapted for a field of view greater than 180 degrees. Real omnidirectional cameras project points in front of the camera to one point and points behind the camera to a different point. The representation of image points by lines assigns to a single image point $\mathbf{p}$ points on a ray passing trough an optical center that are in front as well as behind of the camera. Thus, all points from the ray project to the same point in the projection plane. It allows to represent only scene points lying in a halfspace including an image plane as a border. So, the perspective model is sufficient for directional cameras which cannot see two halfspaces divided by the plane containing the optical center at the same time, but it cannot be used for omnidirectional cameras.

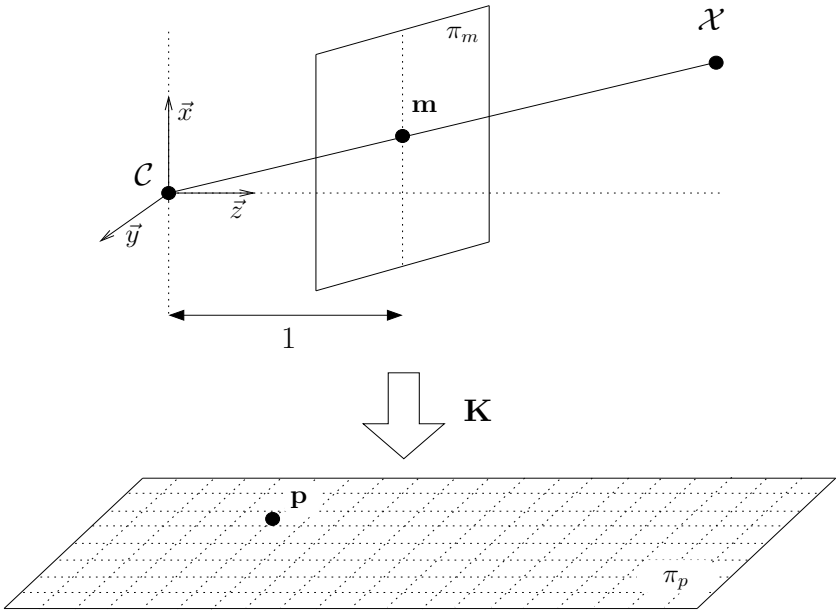


Figure 2.1: *Planar perspective projection.*

2.3.2 Spherical perspective projection

The spherical perspective projection model maps all scene points $\mathbf{X}$ on the unit sphere $\mathbb{S}^2 = \{\mathbf{X}_s \in \mathbb{R}^3 | \|\mathbf{X}_s\| = 1\}$. The constraint that the scene points are in front of the camera can be imposed even for a field of view greater than 180 degrees. The scale factor λ relating points on the sphere to the 3D points must be positive: $\exists \mathbf{X}_s \in \mathbb{S}^2 \implies \exists \lambda > 0 | \mathbf{X} = \lambda \mathbf{X}_s$.

From the unit sphere, we can then apply the projection function noted Π that depends on the intrinsic parameters of the sensor : $\Pi : \Upsilon \subsetneq \mathbb{S}^2 \rightarrow \Omega \subset \mathbb{R}^2$. Π is not defined on *all* of $\mathbb{S}^2$ because we wish Π to be bijective which cannot be the case between $\mathbb{S}^2$ and $\mathbb{R}^2$ as they do not share the same topology. If Π is bijective, Π^{-1} will relate points from the image plane to their projective rays.

2.3.3 Omnidirectional projection

Recalling the standard projection model $\mathbf{p} = \mathbf{P}\mathbf{X}$. Where $\mathbf{p}$ is the projection of the world point $\mathbf{X}$, $\mathbf{P}$ is the projection matrix which comprises the intrinsic parameters and the projection of the world point to a normalised plane [Fau93, HZ00]. For omnidirectional cameras one has: $\mathbf{p} = \mathbf{P}\mathcal{F}(\mathbf{X})$. Where $\mathcal{F}$ is the function introduced by the reflection of the light rays at the mirror. This function depends on the mirror shape and is in general non-linear. Stated simply it determines the point of the mirror where the ray coming from the 3D point $\mathbf{X}$ is reflected and directed towards the camera's optical center.

Depending on the particular camera and mirror setup, the light rays incident to the mirror surface may all intersect at a virtual point. In this case the system is re-

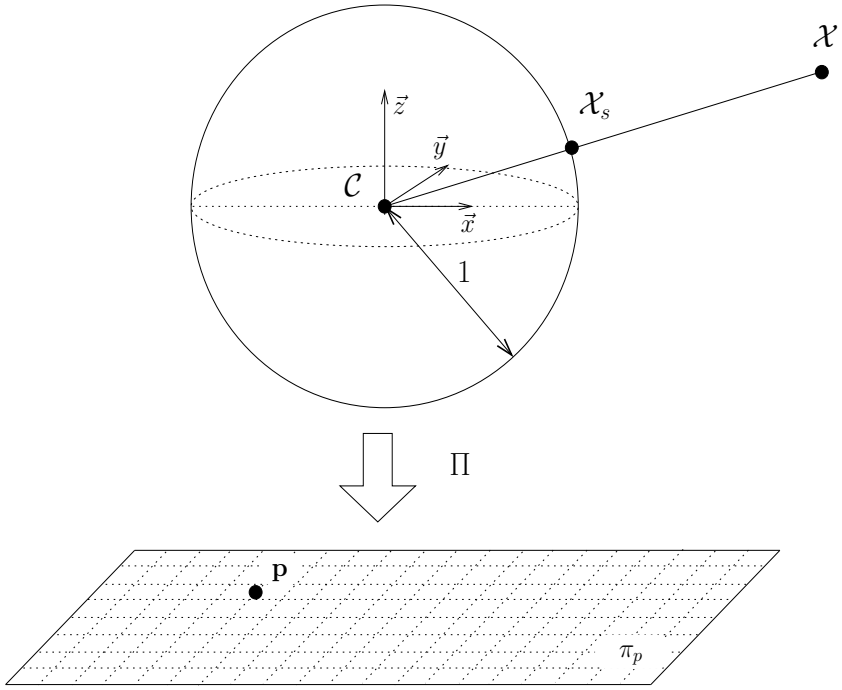


Figure 2.2: *Spherical perspective projection.*

ferred as having a *Single Projection Centre*. A single projection centre is a desirable property as it permits the generation of geometrically correct perspective images from the pictures captured by the omnidirectional camera. This is possible because, under the single view point constraint, every pixel in the sensed image measures the irradiance of the light passing through the viewpoint in one particular direction. When the geometry of the omnidirectional camera is known, that is, when the camera is calibrated, one can precompute this direction for each pixel. Therefore, the irradiance value measured by each pixel can be mapped onto a plane at any distance from the viewpoint to form a planar perspective image.

Baker and Nayar derived the complete class of central catadioptric sensors that have a single viewpoint under the assumption of the pinhole camera model. They identified four possible configurations of camera and mirror that are non degenerate and two configurations that are degenerate. Figure 2.3 depicts the four non degenerate configurations combined an ortographic camera with a parabolic mirror or a perspective camera associated to a hyperbolic, elliptical or planar mirror. The degenerate configurations (spherical mirror and conical mirror) cannot be used to construct cameras with a single effective view point. Indeed, a sphere can be seen as the limit of an ellipse when the two focal points concide; thus, to obtain the single viewpoint property, we would need to place the camera in the center of the sphere. We would then only see the camera itself. By putting the camera in another position, we obtain a caustic. Similarly, a conical mirror is an interesting example of the limit of the pinhole camera model. The single viewpoint constraint imposes that the cone be situated in front of the camera with the vertex at the focal point.

For a pinhole camera, this would mean no light reaches the imager.

2.3.4 Unified Projection Model

A unifying theory for all central catadioptric systems was proposed by Geyer and Daniilidis [GD00, Gey03]. They showed and proved that every catadioptric (parabolic, hyperbolic, elliptical) and standard perspective projection is isomorphic to a projective mapping from a sphere (centered in the effective viewpoint) to a plane with the projection center on the perpendicular to the plane. Figure 2.3 shows the entire class of central catadioptric sensors. Table 2.1 details the equations of the 3D surfaces and the relation between the standard (a, b) parameterisation and the (p, d) parameters used in Barreto's model [Bar03].

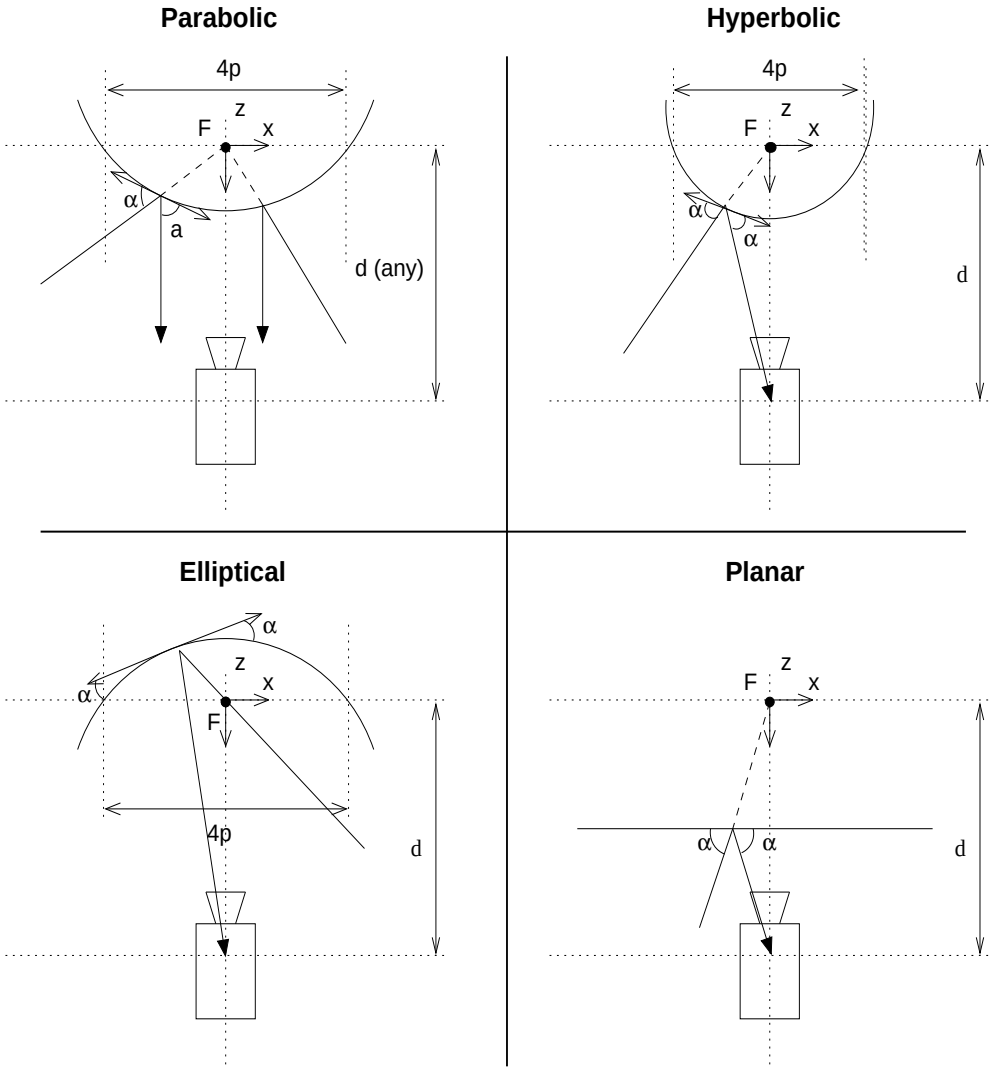


Figure 2.3: *Classes of central catadioptric sensors with a single viewpoint.*

Barreto introduced a modified version of this unifying model [Bar03]. The right

side of Figure 2.4 depicts this model. The mapping between points in the 3D world points and points in the catadioptric image plane is split into three steps. The left side of Figure 2.4 shows this mapping. World points are mapped into an oriented projective plane by a linear function described by a 3×4 matrix (similar to the projective camera model refered in [HZ00]). The oriented projective plane is then transformed by a non-linear function $\mathbf{h}()$. The last step is a collineation in the plane depending on the mirror parameters, the pose of the camera is relative to the reflective surface and the camera intrinsic parameters. Matrix $\mathbf{M_c}$ and function $\mathbf{h}()$ are respectively provided by (2.2) and (2.3). Matrix $\mathbf{M_c}$ only depends on the mirror type and shape. Parameters ξ and φ appear in Table 3.1. $\mathbf{K_c}$ is the matrix of camera intrinsic parameters. Matrix $\mathbf{R_c}$ is a 3×3 rotation matrix which models the possible rotations of the camera with respect to the coordinate system of the camera (See Figure 2.5).

The model obtained is general, intuitive and isolates the non-linear characteristics of general catadioptric image formation in a single function $\mathbf{h}()$. Function $\mathbf{h}()$ transforms oriented projective points $\mathbf{x}$ in sensor coordinates into points $\overline{\mathbf{x}}$. This non-linear transformation, presented in (2.3), has the following interpretation. Without loss of generality, consider that the world and the sensor coordinates system are the same, with origin $\mathbf{O}$ in the effective viewpoint, and an unit sphere centered in $\mathbf{O}$ (See Figure 2.4). To each visible scene point $\mathbf{X_h}$ corresponds an oriented projective ray $\mathbf{x}$ joining the 3D point with the effective projection center. The projective ray intersects the unit sphere in a single point $\mathbf{X_c}$. Consider a point $\mathbf{O_c}$ with coordinates $(0, 0, -\xi)^\top$ in the sensor coordinates system. To each $\mathbf{x}$ corresponds an oriented projective ray $\overline{\mathbf{x}}$ joining $\mathbf{O_c}$ with the intersection point $\mathbf{X_c}$ in the sphere surface. The non-linear mapping $\mathbf{h}$ corresponds to projecting the scene in the unit sphere surface and then re-projecting the points on the sphere into a plane from a novel projection center $\mathbf{O_c}$. Points in the catadioptric image plane $\widehat{\mathbf{x}}$ are obtained after a collineation $\mathbf{H_c}$ of the 2D projective points $\overline{\mathbf{x}}$.

The novel projection center $\mathbf{O_c} = (0, 0, -\xi)^\top$ only depends on the mirror parameter ξ. For a parabolic mirror $\xi = 1$ and $\mathbf{O_c}$ belongs to the unit sphere surface. For hyperbolic and elliptical case $\mathbf{O_c}$ is inside the unit sphere along the negative z-axis. The planar mirror is a degenerate case of central catadioptric projection where $\xi = 0$ and $\mathbf{O_c}$ is coincident with $\mathbf{O}$. We thus fall back to the standard projection model with an extra normalization to the sphere.

$$\mathbf{M_c} = \begin{pmatrix} \varphi - \xi & 0 & 0 \\ 0 & \xi - \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2.2}$$

$$\mathbf{h}(\mathbf{x}) = \begin{pmatrix} \frac{x}{\sqrt{x^2+y^2+z^2}} \\ \frac{y}{\sqrt{x^2+y^2+z^2}} \\ \frac{z}{\sqrt{x^2+y^2+z^2}} + \xi \end{pmatrix} \tag{2.3}$$

Table 2.1: Conic equations

Parabola	$\sqrt{x^2+y^2+z^2} = 2p - z$
Hyperbola	$\frac{(z-\frac{d}{2})^2}{a^2} - \frac{x^2}{b^2} - \frac{y^2}{b^2} = 1$
Ellipse	$\frac{(z-\frac{d}{2})^2}{a^2} + \frac{x^2}{b^2} + \frac{y^2}{b^2} = 1$
Plane	$z = \frac{d}{2}$

$a = \mp 1/2(\sqrt{d^2+4p^2} \pm 2p)$. Negative for a hyperbola. Positive for an ellipse.

$b = \mp\sqrt{p(\sqrt{d^2+4p^2} \pm 2p)}$. Negative for a hyperbola. Positive for an ellipse.

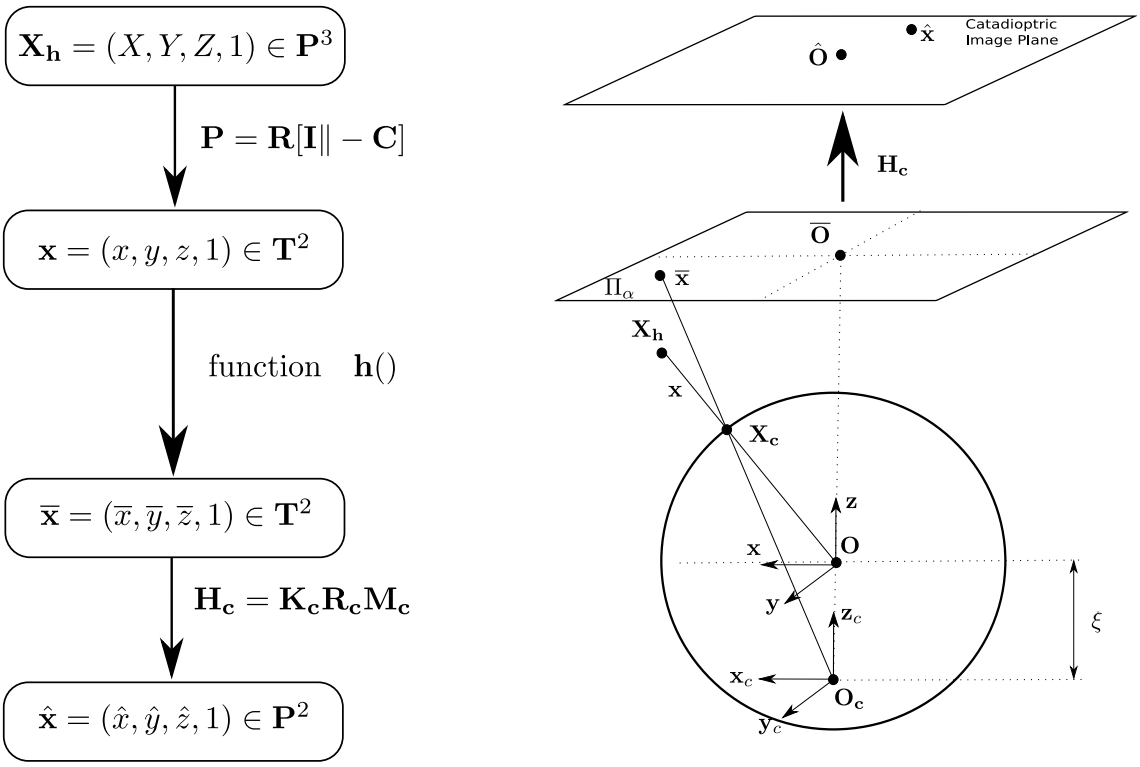


Figure 2.4: *Left. Steps of Barreto's projection model. Right. Barreto's unifified projection model for central catadioptric image formation.*

Recently, this projection model has been very useful in central catadioptric camera applications due to its simplicity. This thesis work is not an exception. A slightly modified version of this projection model will be used throughout this work. The difference from the model of Barreto, is the use of a generalised focal length that depends on the focal length of the camera and on the mirror shape; so that we consider the sensor to be a single imaging device and not the combination of a camera and a mirror. We will present in detail the projection model that will be used in this work in the Section 3.2.1 *Warping Function* of the next chapter.

Table 2.2: Unified model parameters

	ξ	φ
Parabolic	1	$1+2p$
Hyperbolic	$\frac{d}{\sqrt{d^2+4p^2}}$	$\frac{d+2p}{\sqrt{d^2+4p^2}}$
Elliptical	$\frac{d}{\sqrt{d^2+4p^2}}$	$\frac{d-2p}{\sqrt{d^2+4p^2}}$
Planar	0	1
d: distance between focal points $4p$: latus rectum		

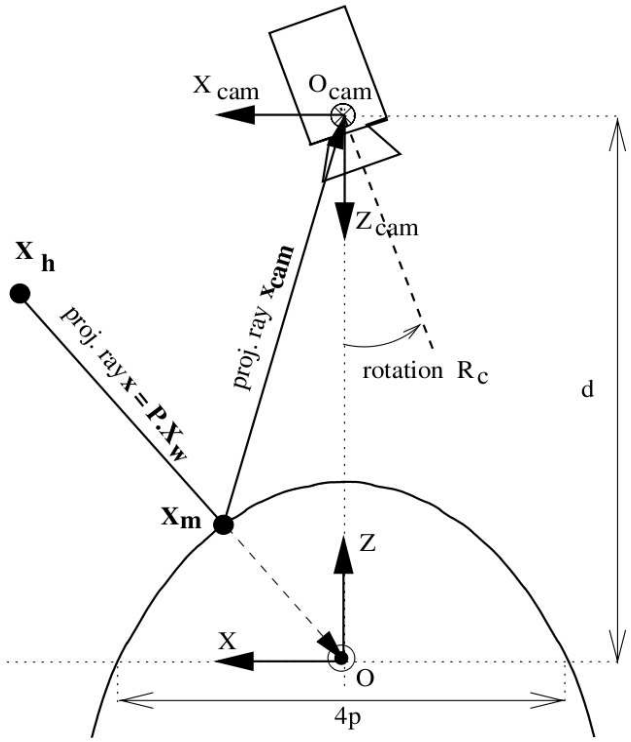


Figure 2.5: *Image formation process. Hyperbolic situation (courtesy of Barreto [Bar03]).*

2.4 Classical Calibration

Camera calibration is the process of modeling the mapping between 3D objects and their 2D images. This process if often required when recovering 3D information from 2D images. More precisely, we have to know the intrinsic parameters and extrinsic parameters.

The *intrinsic* parameters describe the imaging geometric characteristics of the camera. They depend of the used projection model. Among the intrinsic parameters

that most approaches estimate we find :

- The focal length, f.
- The scaling factors for row pixels and column pixels, k_u and k_v.
- The skew factor, s.
- The aspect ratio, $r = \frac{k_u}{k_v}$.
- The principal point, (u_0, v_0).
- The tangential and radial distortion parameters [1].
- The mirror shape parameter(s).

The *extrinsic* parameters represent the orientation and position of the camera with respect to the world coordinate system. These paremeters include a rotation matrix $\mathbf{R}$ between the coordinate system from 3D world coordinates to 3D camera coordinates and a translation vector $\mathbf{t}$ between the coordinate system from 3D world coordinates to 3D camera coordinates.

Classical calibration techniques, as opposed to self-calibration techniques, proceed by analyzing an image of one or several reference objects whose geometry is accurately known. They work in two steps:

1. Some features are extracted from the images by means of standard image analysis techniques. These features are generally points, lines or conics.
2. The features are given as input to an optimization process which searches for the projection parameters that best project the three-dimensional model onto them.

In what follows, we will give a brief overview of the different methotds that have been developed classically.

2.4.1 Pattern Calibration techniques

Methods based on point features lying on a calibration grid (normally a checkboard) require the user to take one picture of the pattern shown covering most of the image [SB08, BPS+08, CBDD00] or several pictures of the pattern shown at a few different positions and orientations around the mirror [MBMR06, SMS06, DWW07, GSB09].

[1] Nonlinear intrinsic parameters such as lens distortion are also important although they cannot be included in the linear camera model described by the intrinsic parameter matrix. Modern camera calibration algorithms estimate these intrinsic parameters as well.

Then, the corner points of the checkboard are detected manually. Figure 2.6 shows sample images of a checkerboard used to calibrate an omnidirectional camera.

Corner extraction for calibration grids, despite being trivial for perspective cameras, is still problematic for images with strong non-linear distortions. In the perspective case we typically indicate the area of interest by manually clicking 4 corners; they enable the estimation of a homography and the projection of the grid onto the image. The final positions of all corners are accurately determined by refining the initial estimate using image processing techniques. Such a procedure has not been possible until now for non-conventional imagery with non-perspective distortions.

In [MBMR06], like in the perspective case, the user is asked to click on the four grid corners of each image, but some steps (considering drawbacks) must be made before in order to make a first calibration approach of the camera's intrinsic parameters. Among these steps we can find : first, the user is asked to say what shape of mirror was used to get the images; second, the user must click in the center and the border of the image mirror. Finally, the user must click three no-radial points in a given pattern image. Although this calibration method has some drawbacks, accurate calibration for all the variety of central catadioptric cameras and fish-eye lenses can be done because the authors adopted the well-known sphere camera model to describe the catadioptric projection.

The proposed method in [SMS06], unlike [MBMR06] only needs the intervention of the user when its automatic corner detection is not capable to detect all the corners due to bluring or heavy distortions in the images. The user must click the remaining corners. Since the mirror shape is described by a polynomial, this algorithm stays valid for all central catadioptric cameras and fish-eye lenses. Both methods, [MBMR06] and [SMS06], compute a least square linear minimization method to find the intrinsic and extrinsic parameters.

In [GSB09] calibration for all central catadioptric cameras is feasible. Like in [MBMR06] the authors adopted the sphere camera model to describe the catadioptric projection. The authors showed that, using the so-called lifted coordinates, a linear relation mapping the grid points to the corresponding points on the image plane can be written as a 6×6 matrix $\mathbf{H}_{cata}$, which acts like the classical 3×3 homography for perspective cameras. They showed how to compute the image of the absolute conic (IAC) from at least 3 homographies and how to recover from it the intrinsic parameters of the catadioptric camera. Altough the algorithm considers a linear relation, at least three images from a checkboard pattern with more than 48 corners are necessary and their corners have to be manually detected.

In [SB08] the authors proposed a work similar to [GSB09]. They showed that in the case of paracatadioptric cameras one homography could be enough for estimating the IAC, thus allowing the calibration from a single image.

In [BPS+08] the authors compute calibration using three perpendicular checkboard

patterns. This work uses the sphere camera model, so it is able to calibrate every single-viewpoint catadioptric system. The algorithm needs a minimun of twenty points distributed in the three different planes to estimate the generic 6×10 projection matrix, which lifts coodinates for the image and the 3D points. The authors show how to decompose this projection matrix to obtain intrinsic and extrinsic parameters. Moreover, they use this estimation of parameters followed by a non-linear optimization to calibrate various types of cameras.

In [CBDD00] the authors use a pattern calibration which is a hollow cube with four interior vertical sides made up of a repetitive square motive. To perform calibration of the conical mirror, the authors select the non-radial points in the image. They characterize all horizontal straight lines of the pattern calibration, projecting them in the picture under the shape of the curve. Next, the authors approximate with elliptic functions all curves in the picture. Then, radial straight line intersections with these curves are calculated. These intersections give the set of 2D calibration points. The set of 2D and 3D calibration points give an overdetermined system. A non-linear minimization method is utilised to solve this system.

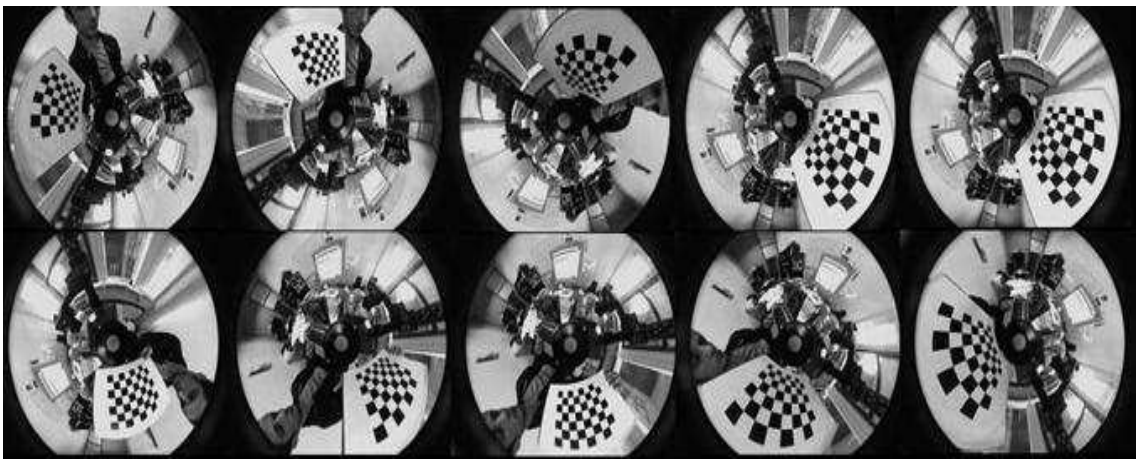


Figure 2.6: *Sample images of a checkerboard pattern used to calibrate an omnidirectional camera.*

2.4.2 Line based calibration

Methods based on lines calibrate omnidirectional cameras using geometric properties of the conic. They detect lines either automatic or manually, thus calibration is performed. Figure 2.7 depicts detections of lines in omnidirectional images for calibration purposes.

Geyer and Daniilidis [GD02] calibrated para-catadioptric cameras from at least three line images. The authors showed that vanishing points lie on a conic section which encodes the entire calibration information. Thus, projections of two sets of parallel lines suffice for intrinsic calibration. Howewer, this property does not apply to non-

parabolic mirros. Therefore, the proposed technique cannot be easily generalized to other kinds of sensors.

In [BA05] Barreto and Araujo studied the geometric properties of line images under the central catadioptric model, and gave a calibration method for all kinds of central catadioptric cameras. Both methods [GD02, BA05] involve conic fitting, therefore, errors can occur when only a partial contour is available.

Wu *et al.* [WLH06] introduced a shift from the central catadioptric model to the pinhole model. The constraint that image points belonging to a space line is collinear is used to calibrate paracatadioptric-like cameras. This method does not need fitting of partially visible conic and is totally linear.

In [YZ05] the method starts by quantizing the intrinsic parameters in their value ranges and then, for each calibration set, a line detection algorithm is performed in sphere space. The best parameters are those maximizing a cost function defined as the number of pixels that belong to the 3 most dominant lines. One of the disadvantages is that the search space is analyzed exhaustively whereas it is extremely huge (5 dimensions of real values). Moreover, the line extraction algorithm is based on the Hough transform which is known to be slow and sensitive to parameters sampling.

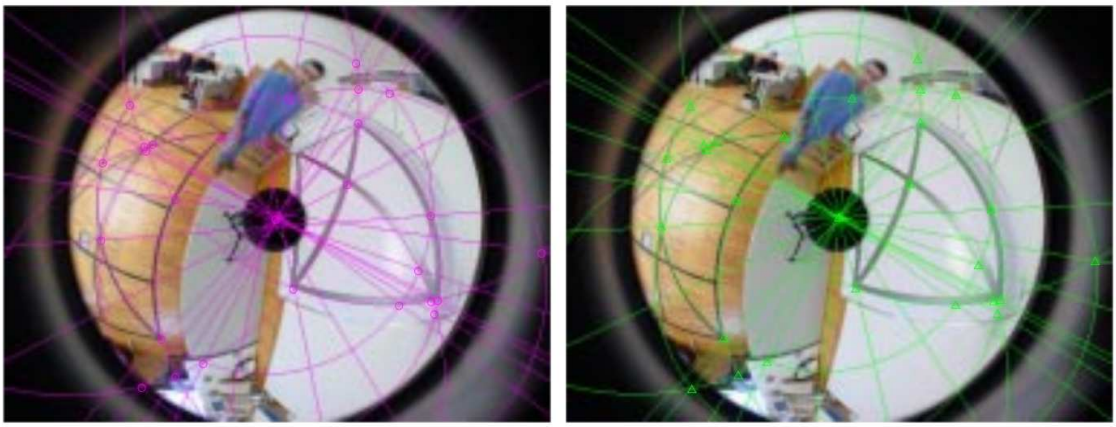


Figure 2.7: *Detection of lines for calibration purposes.*

2.4.3 Points based calibration

Calibration techniques based on points exploit known 3D space points to calibrate the sensor. Figure 2.8 shows a calibration system where the 3D positions of points in the enviroment are well known.

In [Ali01] Aliaga proposed an approach to estimate camera intrinsic and extrinsic parameters for a paraboloidal catadioptric system, where the mirror center was manually determined. Their model relaxes the requirement of a perfect orthographic projection, integrates radial distortion correction, and compensates for minor mirror misalignment.

Vasseur and Mouaddib [Vas04] calculated intrinsic parameters by a non-linear method with 3D space points. The authors placed the catadioptric system in an opened cube with a grid of points on each side. In this way, the points of the pattern were distributed over the whole catadioptric image. The pattern contained 112 points and the authors calculated six extrinsic parameters as well as the intrinsic parameters of the camera. The optimization was performed by the minimization of the quadratic error between the selected points and those computed by the model.

Wu and Hu [WH05] introduced the invariants of 1D, 2D and 3D space points and then used them to compute the camera's principal point with a quasi-linear method.

Figure 2.8: *Point based calibration system. The 3D positions of the points in the calibration pattern are well known (P. Vasseur [Vas04]).*

2.4.3.1 Sphere based calibration

Sphere based calibration techniques detect sphere shapes and then their contours are exploited to calibrate the sensor. Figure 2.9 shows an omnidirectional image with a sphere for calibration purposes.

Ying and Hu [YH03] analyzed the relation of the camera intrinsic parameters and the sphere imaged contour. Then, they applied the relation to calibrating central catadioptric cameras. The minimum number of sphere images (three) for the calibration of different types of central catadioptric cameras was clarified in this paper. In order to efficiently solve the non-linear equations of the intrinsic parameters, the authors presented a two-stage calibration technique which divided the intrinsic parameters into two groups and used the Levenberg-Marquardt algorithm to perform the minimization to estimate the intrinsic parameters. One important observation in this paper is that the methods based on the projection of spheres are more robust and have higher accuracy than those using the projection of lines.

In [XY05] the authors discovered that each sphere image is tangent to the modified image of the absolute conic (MIAC) at two double-contact image points, and a linear calibration method using sphere images is derived from this observation. This algorithm requires three sphere images and recovers all five intrinsic parameters without making assumptions, such as zero-skew or unitary aspect ratio.

Figure 2.9: *Sphere image used for calibration purposes.*

2.4.4 Polarized images calibration

These techniques are based on polorization imaging. Figure 2.10 depicts a polarization system and the needed images to compute the calibration parameters.

Morel and Fofi [MSF07] use the generic calibration concept introduced by Sturm and Ramalingam [SRL06]. The concept considers an image as a collection of pixels, and each pixel measures the light along a particular 3D ray. Thus, calibration is the determination of the coordinates of all pixels' rays. To get the 3D-rays the 3D surface of the mirror has to be computed. The authors compute this 3D surface thanks to the Frankot-Chellappa algorithm. To calibrate the system, the rotating polarizer is placed between the lens and the mirror, and a white paper sheet cylinder is placed around the mirror in order to get unpolarized light. Each light intensity of the pixels is linked to the angle of the polarizer (called α) and the polarization parameters I, Ψ and ρ are computing with at least three images taken with different orientations of the polarizer. To get an automatic calibration of the catadioptric system, a liquid-crystal polarization rotator is used instead of the polarizer. After automatically computing the polarization parameters, by knowing the refractive indexs of the material, the calibration is directly done for every pixel. No image proccesing and no calibration pattern are required.

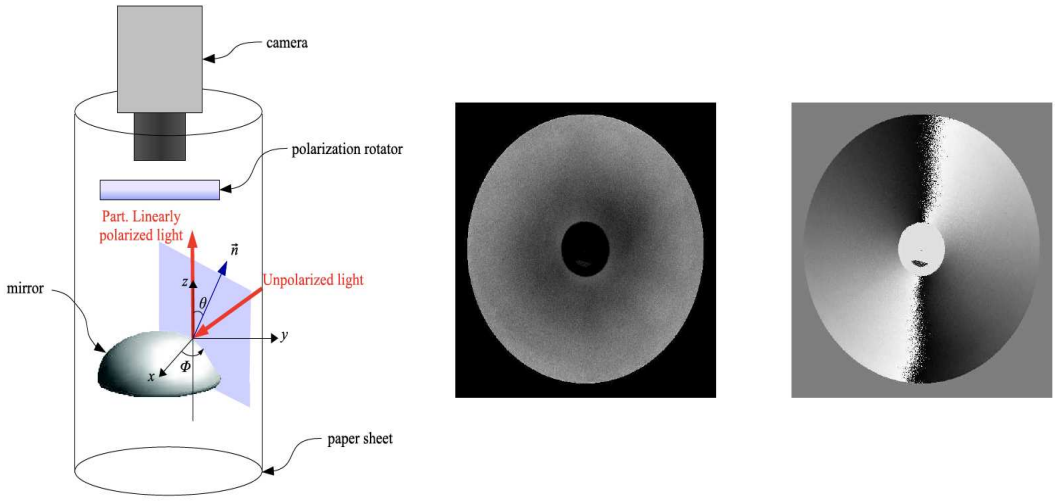


Figure 2.10: *Polarized images calibration system. Left. Polarization system. Middle and right. Images of the polarization parameters that are needed to reconstruct the mirror shape (Morel and Fofi [MSF07]).*

2.4.5 Classical calibration drawbacks

Some remarks can be made for these kind of techniques. All of them differ mainly by the type of mirror taken into account (e.g. hyperbolic and parabolic) and the

considered intrinsic parameters (skewness, aligment, ...) of the projection model. Even though, methods based in a calibration grid are capable of performing accurate calibration for all central catadioptric cameras (provided the calibration pattern is carefully set), corner extraction for calibration grids is still problematic for images with strong non-linear distortions; consequently, manual extraction is often required. It is not possible to calibrate the camera on-line when it is already involved in a visual task (in many applications a calibration pattern is not available). However, even when the camera performs a task, the intrinsic parameters can change intentionally or not (e.g. adjustment of the focal length, mechanical or thermal variations). Methods based on projection of lines have the advantage that a calibration pattern is optional (lines from the real world can be used instead). However, the image of a line can be a line, circle, ellipse, hyperbola or parabola, and in general only a small arc of the conic is visible in the image and it is known that fitting a generic conic when its type is unknown is hard. Therefore traditional fitting methods lead to inaccurate results. The accuracy of the estimated intrinsic parameters depends highly on the accuracy of the extracted conics. Methods based on sphere images permits a more accurate conic fitting since the whole sphere contour is visible. However, it requires the presence of spheres in the scene which cannot be applied in real world applicactions. Finally, methods based on polarized images work for any shape or mirror (regular or not) but errors can occur in the top of the mirror where the normals are oriented near the optical axis. Morever, a calibration pattern is not required but a complex polarized system is required.

In chapter 4, we will detail how to self-calibrate omnidirectional cameras on-line based on the image intensity. Before achieving this objective, we will show in the following chapter how to efficiently track a plane in an omnidirectional image without requiring the prior calibration of the sensor. This will give us the basis for self-calibrating omndirectional cameras on-line by using several of the tracked views.

Chapter 3

Uncalibrated Visual Tracking

In this chapter we will show how to efficiently track a plane in an omnidirectional image without requiring the prior calibration of the sensor. We will start by giving an overview of tracking approaches to set the context for our work on omnidirectional plane-based visual tracking. We will then concentrate more specifically on sum of squared differences (SSD) tracking. We will explain why SSD tracking is particularly well suited for robotics tasks. Then, we will explain in detail the uncalibrated visual tracking algorithm. Finally, we will present experimental results. These results will show that our method is able to track planar objects with an uncalibrated catadioptric camera and thus can be helpful in robotic applications where camera calibration is impossible or hard to obtain.

3.1 Visual Tracking

Visual tracking is a critical task in many computer vision applications such as surveillance, robotics, human computer interaction, vehicle tracking, medical imaging, etc. The challenges in designing a robust visual tracking algorithm come from the presence of noise, occlusion, varying viewpoints and illumination changes. A variety of tracking algorithms have been proposed to overcome these difficulties.

Although it is not always obvious how to characterise an approach, visual tracking approaches can be classified according to several properties as:

- 3D model-based/model-free tracking.
- Matching-based/direct tracking.
- 2D/3D tracking.

In 3D model-based/model-free tracking techniques, we can either assume that the tracked object can be parameterised by a certain surface or structure (planes, quadratics, cubes, ...), or we characterise it only by its properties (color, texture, rigid or deformable, ...).

In matching-based/direct tracking techniques, we can either extract features in the image and then look for similar features in the incoming image without using prior knowledge of the camera motion, or the approach assumes small displacements between frames and processes the image information directly.

In 2D/3D tracking techniques, we can either track an object in the image or estimate the 3D motion of the object in the scene from its image.

In real-time applications, the main requirements of a tracking algorithm are efficiency, accuracy and robustness. Visual tracking techniques such as tracking by matching and recursive tracking achieve these requirements. Lets discuss them in order to set the context for our work on omnidirectional plane-based visual tracking.

3.1.1 Tracking by matching

Tracking by matching techniques work in three steps. First, features are extracted in the first image. Second, features are extracted in the second image and third, features are associated in the two images through a distance measure.

These techniques do not search for the points extracted in the first image directly in the second image because it is supposed to obtain a gain in terms of computation and robustness. For example, if we use Harris points [HZ00], we only need to correlate between features instead of searching in the whole image. Furthermore there is a smaller chance of having bad associations since we only correlate between features that we expect to be able to associate. Of course, if there are a lot of features and that the distance measure is not very discriminating, we can expect to have many outliers. Matching in the whole image also has the disadvantage of being computationally expensive but has the advantage of making it possible to match over large distances. The efficiency and robustness of the data association can be enforced using a priori knowledge of the motion of the camera or of the tracked object.

Outlier rejection is an essential process in these type of methods. For example, if we are estimating the motion of the camera from feature points, we might use the epipolar constraint (through the essential matrix or fundamental matrix) to remove outliers. The constraint is imposed by the properties of projective geometry. It is also possible to add constraints such as planarity (planar homography matrix) or even search for specific objects.

The outlier rejection process and the salient feature extraction can make it difficult to obtain frame-rate tracking. RANSAC [FB81] is often used but it is time-consuming. Improvements have been made to the standard RANSAC approach [MC02] and modern computers make it possible to obtain high frame rate computation even with these processing steps. Furthermore, SIFT points [Low04] and learning-based techniques [LLF05] produce fewer outliers, the downside being a higher computational burden compared to Harris points. However fewer outliers can remove altogether the need for RANSAC. Faster robust estimators (Tukey, Huber) are then sufficient. The advantages of these approaches come from the robustness to occlusion and to change in intensity, and also from the possibility of tracking objects with large displacements in the image.

These methods have been applied successfully for tracking planes [SFZ00], head pose [TMdCM02] or deformable objects [PLF05].

3.1.2 Direct tracking

Recursive tracking algorithms between two images work in four steps as follows :

1. features are extracted in the first image,
2. the features are projected in the second image,
3. a distance measure between the information in the second image and the features from the first image is applied and
4. a direction that will minimise the error is found.

These tracking techniques assume that the motion of the features in the image between the subsequent frames is small and that the objects are considered as Lambertian[1]. This assumption is reasonable if the camera frame rate is high compared to the object motion and if the computation of steps 2, 3 and 4 hold within the frame rate.

The approach is however generally more sensitive to occlusion and less well adapted to fast motion.

One of the first recursive tracking techniques was an edge-based tracking. An edge is extracted in an image and reprojected in the following image. The distance to minimise is then typically obtained by searching edge points along the normals to the initial edge. The advantage of edge-tracking is its robustness to changes in intensity but it is sensitive to occlusion. The minimisation can either be image-based or 3D (we minimise the 3D pose of the object according to the projection of its edges in the image).

Sum of squared differences (SSD) tracking can be traced back to the work by Lucas and Kanade [LK81] and later Tomasi [ST94b]. SSD measures the difference in intensity between a portion of the first image reprojected in the second image. The minimisation (based on the image gradient) can be imaged-based (2D) for example searching for the translation that gives the smallest reprojection error. It can also be 3D or model-based by reprojecting a 3D object and minimising the difference in the image over the position (rotation and translation). The advantage of this approach is precision (all the information is being used) and speed. This is why these techniques are particularly well adapted to robotic tasks such as motion estimation and vision-based robot control. Compared to matching approaches, SSD tracking is generally faster and more precise. The downside is the need for a strong overlap between the reprojected and the real object for the system to converge.

[1] A surface exhibits Lambertian reflectance if light falling on it is scattered such that the apparent brightness of the surface to an observer is the same regardless of the observer's angle of view. More technically, the surface luminance is isotropic. For example, unfinished wood exhibits roughly Lambertian reflectance, but wood finished with a glossy coat of polyurethane does not, since specular highlights may appear at different locations on the surface.

The tracking algorithm that we will describe in this Chapter is a plane-based 3D SSD visual tracking using an uncalibrated omnidirectional sensor. A brief overview of the different approaches found in the literature will be described next.

3.1.2.1 Overview of SSD Visual Tracking with Omnidirectional Sensors

Visual tracking with omndirectional sensors is not a straightforward task. The non-linear projection model resulting in changes of shape in the image makes the direct use of methods such as KLT [LK81, ST94b] nearly impossible. Therefore, in the literature, only a few works on SSD tracking for omnidirectional sensors have been developed [MBMR06, BMH02, JO02].

Let us describe these approaches a bit and show how parametric models such as the homography-based approach presented in [MBMR06, TG04, HAMAM05] as well as the one described in this chapter are well adapted to this problem.

In [BMH02], Barreto *et al.* proposed a method for tracking omnidirectional lines using a contour-to-point tracker to avoid the problem of quadric-based catadioptric line fitting (See Figure 3.1). The authors have shown how the method can be adapted for every kind of catadioptric omndirectional sensor using the unified projection model on the sphere.

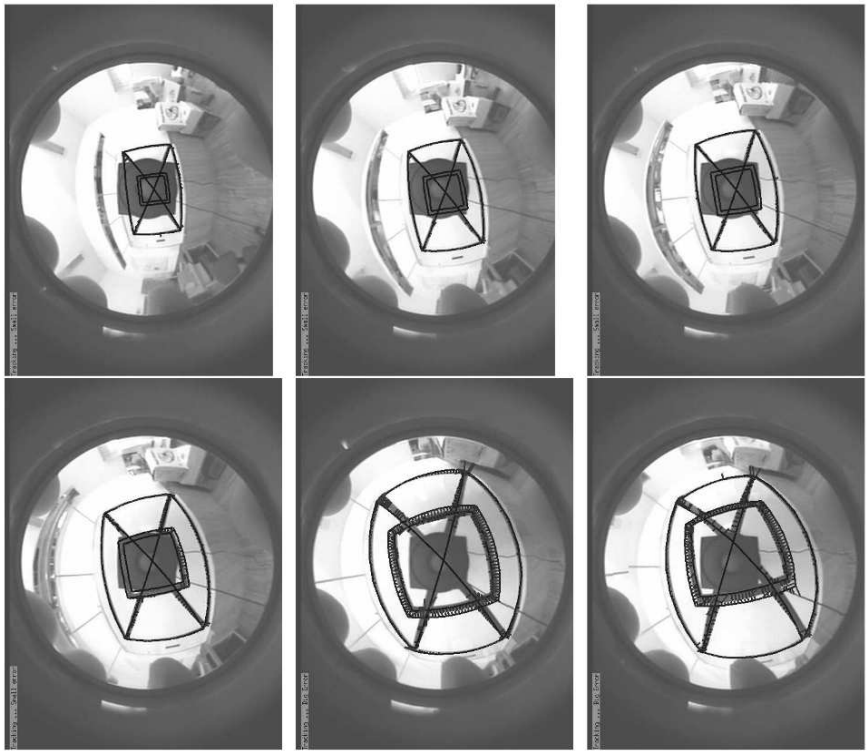


Figure 3.1: *A tracking sequence. The rectangular object moves towards the camera and in a direction perpendicular to the camera (Courtesy of Barreto [BMH02]).*

In [JO02], Jun Okamoto and Valdir Grassi proposed a method for tracking a moving target on an omnidirectional vision system (See Figure 3.2). This algorithm assumes that the deformation suffered by the target image region at each incoming image can be modelled by an affine transformation. Since the affine transformation model does not hold for the acquired omnidirectional image, the tracking algorithm does not work on the omnidirectional image directly. Morever, the method works only for hyperbolic mirrors.

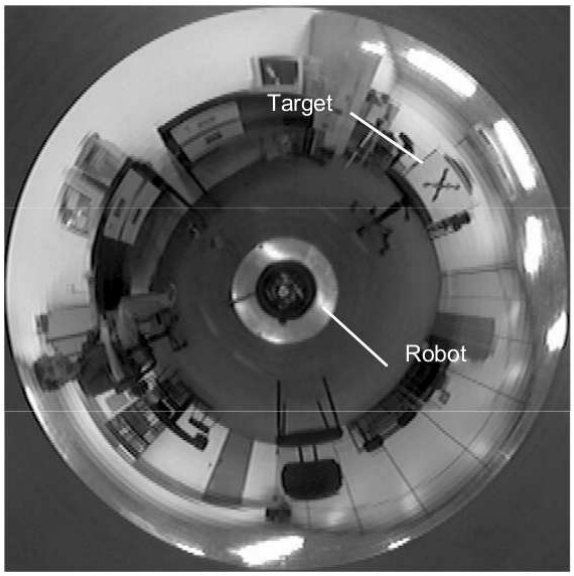


Figure 3.2: *A tracking target. Moving target in an omnidirectional image (Courtesy of Jun Okamoto [JO02]).*

In [MBMR08], which is the closest work to our approach, C. Mei *et al.* proposed a direct method for tracking piecewise planar objects with any central catadioptric camera. The approach deals with a nonlinear optimization problem based on the image intensity. It assumes that the motion of the features in the images from subsequent frames is small. The objects are considered as Lambertian. This assumption is reasonable if the camera frame rate is high compared to the object motion and if the computation of steps two to four hold within the frame rate. The authors extended the standard notion of homography to omnidirectional cameras using the unified projection model on the sphere. With this approach, the difficulty of tracking with these devices due to the non-linear projection model was completely avoided (See Figure 3.3).

We want to point out, even if it has not been mentioned in the overview, that the existing algorithms need a precisely calibrated omnidirectional sensor. Here lies the importance of our approach. The tracking algorithm that we will present does not need prior knowledge of the sensor parameters.

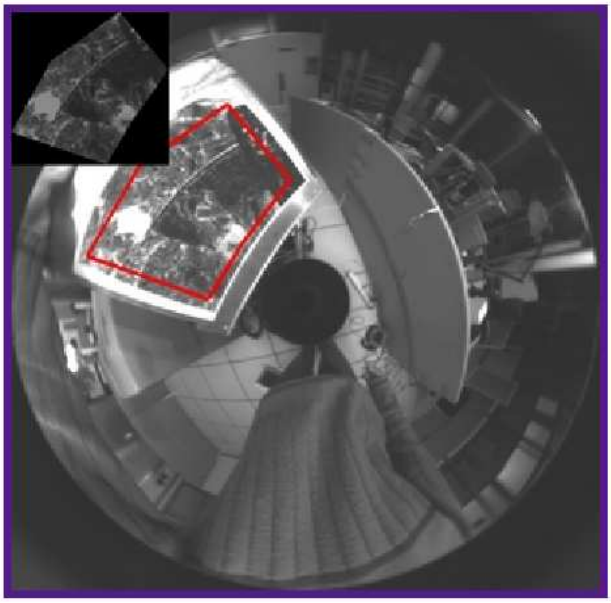


Figure 3.3: *Tracking of a plane. Moving plane in an omnidirectional image (Courtesy of C. Mei [MBMR06]).*

Our approach is based on [MBMR08]. But let us remark that the optimization problem we address is much more challenging than the problem in [MBMR08] because [MBMR08] uses a calibrated camera and therefore the intrinsic parameters are known. We have to also estimate the intrinsic parameters which makes our problem much more challenging.

3.2 Uncalibrated Visual Tracking Algorithm

The visual tracking algorithm will essentially be an image registration problem which will be related directly to the grey-level brightness measurements in the catadioptric images. Figure 3.4 illustrates an example for a monocular camera. Let I_0 be the catadioptric reference image and a region of size $\mathcal{R}$ of I_0 be the *reference plane*. This region of size $\mathcal{R}$ corresponds to the projection of a 3D planar region of the scene. Let all possible unknown parameters be collected together in $\mathbf{x}$. Once we find the optimal parameters $\overline{\mathbf{x}}_1$ to align the image I_1 with the reference template $\mathcal{R}$ we look for the incremental parameters $\mathbf{x}$ to align the image I_2 with $\mathcal{R}$, and so on. If we are able to converge at each step, we obtain the optimal parameters estimation between the first and the last view without drift.

It is well understood that, throughout the sequence, the reference plane would eventually need to be updated (e.g. when the overlap between the reprojected object and the reference object is small).

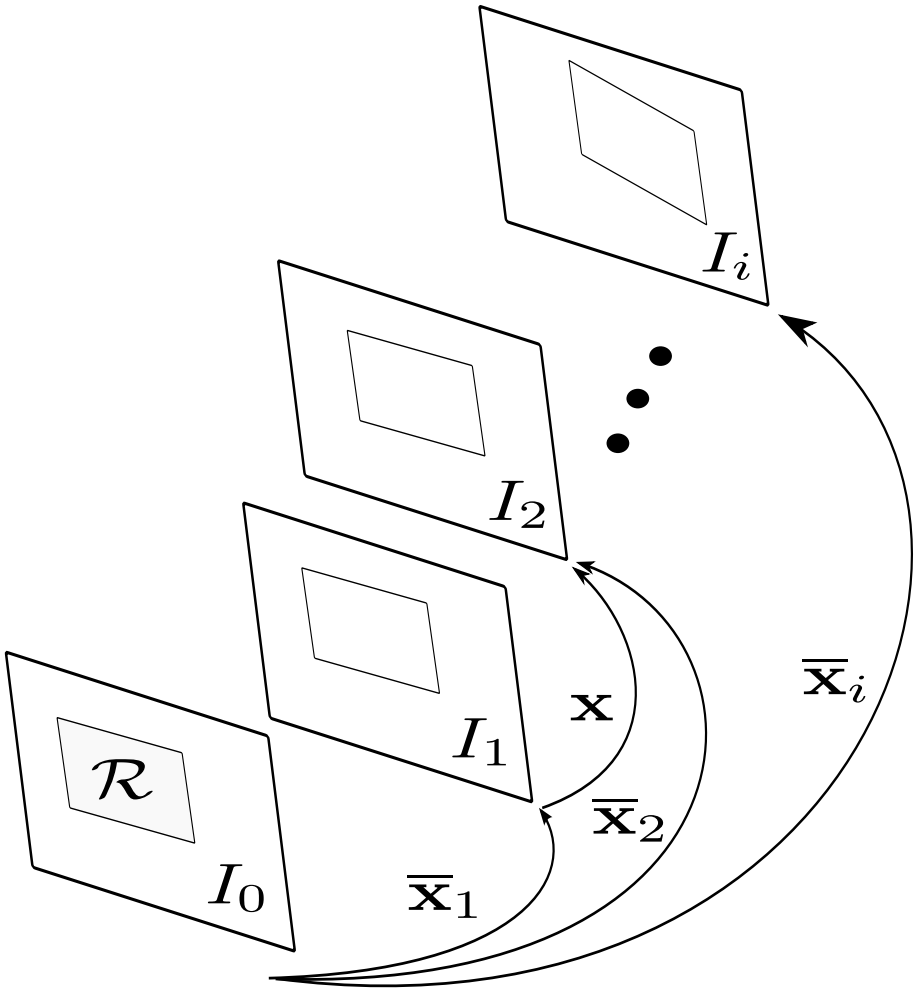


Figure 3.4: *Incremental image registration.*

3.2.1 Warping Function

The function which relates the image points $\mathbf{p}$ belonging to a planar region in a reference image I with the points $\mathbf{p}'$ of this planar region in a current image I' will be called *warping* and will be denoted by $\mathbf{w}$. It depends on a homography $\mathbf{H}$, the mirror parameter ξ and the perspective camera parameters $\mathbf{K}$:

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

The warping function basically includes three transformations (See Figure 3.5) :

1. The transformation between the image plane and the unit sphere.
2. The transformation between spheres.
3. The transformation between the unit sphere and the image plane.

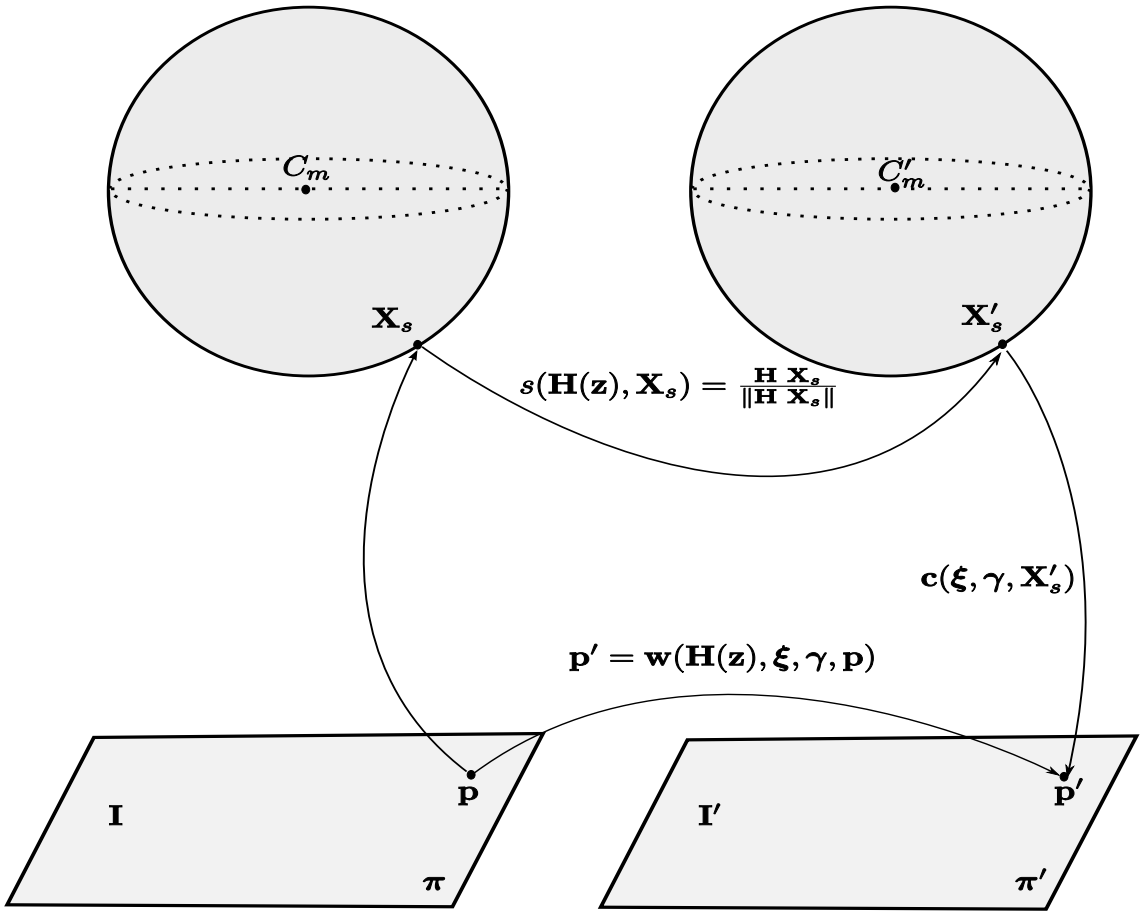


Figure 3.5: *Warping.*

In the following section we will describe these transformations as well as the parameters to be estimated in the tracking algorithm.

3.2.1.1 The transformation between the image plane and the unit sphere

Let $\mathbf{c}^{-1}$ denote the transformation between the image plane and the unit sphere

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

This transformation carries a point $\mathbf{p} = (u\ , v\ , 1)^\top$ measured in the image plane to the unit sphere as follows :

The first step is to apply the inverse projection induced by $\mathbf{K}^{-1}$ to obtain a point $\mathbf{q} = \mathbf{k}^{-1}(\boldsymbol{\gamma}, \mathbf{p}) = \mathbf{K}^{-1}\mathbf{p}$ on the normalized plane. $\mathbf{K}$ is a generalised camera projection matrix. A generalized camera projection matrix indicates that the sensor is considered as a global device instead of a separate camera and mirror. This is particularly important for calibration because it shows that f (the camera focal length) and and η cannot be estimated independently. η is a parameter obtained from mirror parameters of Table 3.1).

Table 3.1: Unified model parameters

	ξ	φ	$\eta = \xi - \phi$
Parabolic	1	$1 + 2p$	$-2p$
Hyperbolic	$\frac{d}{\sqrt{d^2+4p^2}} < 1$	$\frac{d+2p}{\sqrt{d^2+4p^2}}$	$\frac{-2p}{\sqrt{d^2+4p^2}}$
Elliptical	$\frac{d}{\sqrt{d^2+4p^2}} < 1$	$\frac{d-2p}{\sqrt{d^2+4p^2}}$	$\frac{2p}{\sqrt{d^2+4p^2}}$
Planar	0	1	-1
d: distance between focal points $4p$: latus rectum			

$$\mathbf{K} = \begin{pmatrix} k_u f & 0 & u_0 \\ 0 & k_v f & v_0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{M_c} = \begin{pmatrix} k_u \eta & 0 & u_0 \\ 0 & k_v \eta & v_0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \beta_1 & 0 & u_0 \\ 0 & \beta_2 & v_0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.1}$$

k_u and k_v are the scaling factors for row pixels and column pixels. (u_0, v_0) the principal point. The skew factor s is considered zero (a perfect orthogonality of the array camera matrix elements). $\beta_1 = k_u f \eta$ and $\beta_2 = k_v f \eta$ are the generalized focal lengths.

In the function $\mathbf{k}^{-1}(\boldsymbol{\gamma}, \mathbf{p})$, $\boldsymbol{\gamma} = [\beta_1, \beta_2, u_0, v_0]^\top$ contains the camera intrinsic parameters.

The second step projects the point $\mathbf{q} = (x\ , y\ , 1)^\top$, obtained in the previous step, using the inverse function $\alpha = \frac{\xi + \sqrt{(1-\xi^2)(x^2+y^2)}}{x^2+y^2+1}$ proposed by Barreto [BA05] to obtain a point $\mathbf{X}_s = \mathbf{h}^{-1}(\boldsymbol{\xi}, \mathbf{q}) = (\alpha x, \alpha y, \alpha - \xi)^\top$ on the unit sphere. ξ contains the mirror parameter. This parameter defines the shape of the used omnidirectional mirror (See Table 3.1).

The unified projection model has been shown to be valid for fish-eye lenses [CMEM07]. In [CMEM07] the authors have proved that the unified projection model for catadioptric systems can model fish-eye cameras with distortions directly included in its parameters. The validity of this assumption is discussed and compared with other existing models. The authors calibrate their camera using the method designed for central catadioptric cameras proposed in [MR07]. From their analysis one can conclude that a value of ξ bigger than the unity represents a fish-eye camera. Therefore, since we estimate the value of ξ, our method can be useful to track planar objects using a fish-eye camera.

3.2.1.2 The transformation between spheres

Let $\mathbf{s}$ be the function that transforms the points between the spheres

$$\begin{array}{rccc} \mathbf{s}: & \mathbb{SL}(3) \times \mathbb{S}^2 & \longrightarrow & \mathbb{S}^2 \\ & (\mathbf{H}(\mathbf{z}), \mathbf{X}_s) & \longmapsto & \mathbf{X}_s' = \mathbf{s}(\mathbf{H}(\mathbf{z}), \mathbf{X}_s) \end{array}$$

This transformation is performed through an homography. The standard notion of homography[2] can be well adapted for points belonging to the sphere of the unified projection model [MBMR06, HAMAM05]. Therefore, the projection between points $\mathbf{X}_s$ and $\mathbf{X}'_s$, belonging to a planar region of the scene are related by $\mathbf{X}'_s = \mathbf{s}(\mathbf{H}(\mathbf{z}), \mathbf{X}_s) = \frac{\mathbf{H}\mathbf{X}_s}{\|\mathbf{H}\mathbf{X}_s\|}$. $\mathbf{z} = [h_1, h_2, h_3, h_4, h_5, h_6, h_7, h_8, h_9]^\top$ contains the homography parameters. The homography matrix can be expresed as follows:

$$\mathbf{H} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix}$$

A homography is defined up to a scale factor. In order to fix the scale, we choose all the homographies belonging to the special linear subgroup of dimension 3 defined as $\mathbb{SL}(3) = \{\mathbf{H} \in \mathbb{GL}(3) \mid \det(\mathbf{H}) = 1\}$.

3.2.1.3 The transformation between the unit sphere and the image plane

Let $\mathbf{c}$ be the transformation between the sphere and the image plane

$$\begin{array}{rccc} \mathbf{c}: & \mathbb{R} \times \mathbb{R}^5 \times \mathbb{S}^2 & \longrightarrow & \mathbb{P}^2 \\ & (\boldsymbol{\xi}, \boldsymbol{\gamma}, \mathbf{X}'_s) & \longmapsto & \mathbf{p}' = \mathbf{c}(\boldsymbol{\xi}, \boldsymbol{\gamma}, \mathbf{X}'_s) \end{array}$$

This transformation lifts a point $\mathbf{X}'_s = (X'_s, Y'_s, Z'_s)^\top$ on the unit sphere to the image plane as follows :
First, the point $\mathbf{X}'_s$ is changed to a new reference frame centered in $\mathbf{C}_p = (0, 0, -\xi)$. Then, the point is projected to the normalised plane to obtain $\mathbf{q}' = \mathbf{h}(\boldsymbol{\xi}, \mathbf{X}'_s) = \left(\frac{X'_s}{Z'_s - \xi}, \frac{Y'_s}{Z'_s - \xi}, 1\right)^\top$. Finally, the camera projection matrix $\mathbf{K}$ is applied to obtain the point $\mathbf{p}' = (u',\ v',\ 1)^\top = \mathbf{k}(\boldsymbol{\gamma}, \mathbf{q}') = \mathbf{K}\mathbf{q}'$.
If expressions in sections (3.2.1.1), (3.2.1.2) and (3.2.1.3) are combined, the warping function can be written as

$$\mathbf{p}' = \mathbf{w}(\mathbf{H}, \boldsymbol{\xi}, \mathbf{K}, \mathbf{p}) = \mathbf{k}(\boldsymbol{\gamma}, \mathbf{h}(\boldsymbol{\xi}, \mathbf{s}(\mathbf{H}(\mathbf{z}), \mathbf{h}^{-1}(\boldsymbol{\xi}, \mathbf{k}^{-1}(\boldsymbol{\gamma}, \mathbf{p})))))) \tag{3.2}$$

3.2.2 Optimization problem

In this section we will consider that the observed scene is Lambertian and static with respect to the light sources. We will suppose that $\overline{\mathbf{x}} = (\overline{\mathbf{z}}, \overline{\boldsymbol{\xi}}, \overline{\boldsymbol{\gamma}})$ exist such that the current image can be registered with the reference image $I'(\mathbf{w}(\overline{\mathbf{x}}, \mathbf{p}_i)) = I(\mathbf{p}_i)$.

Let us suppose that we have an initial approximation $\widehat{\mathbf{x}}$ of the true parameters $\overline{\mathbf{x}}$. The incremental image registration problem can be formulated as the solution of the following non-linear system :

$$f_i(\widehat{\mathbf{x}} \circ \mathbf{x}(\mathbf{z}, \boldsymbol{\xi}, \boldsymbol{\gamma}), \mathbf{p}_i) = I'(\mathbf{w}(\widehat{\mathbf{x}} \circ \mathbf{x}(\mathbf{z}, \boldsymbol{\xi}, \boldsymbol{\gamma}), \mathbf{p}_i)) - I(\mathbf{p}_i) = \mathbf{0}$$

[2]Two points $\mathbf{X}'$ and $\mathbf{X}$ associated with a homography $\mathbf{H}$ satisfy $\rho'\mathbf{X}' = \rho\mathbf{H}\mathbf{X}$. The standard planar homography matrix $\mathbf{H}$ is defined up to a scale factor: $\mathbf{H} \sim \mathbf{R} + \mathbf{t}\mathbf{n}_d^{*\top}$, where $\mathbf{R} \in \mathbb{SO}(3)$ is the rotation matrix of the camera and $\mathbf{t}$ its translation vector, $\mathbf{n}_d^* = \mathbf{n}^*/d^*$ is the ratio between the normal vector of the plane $\mathbf{n}^*$ and the distance d^* of the plane to the origin of the reference frame.

where $i = \{1, 2, \ldots, \mathcal{R}\}$ ($\mathcal{R}$ is the size of the reference image or template). $\mathbf{p}_i$ represents the coordinates of the pixels of the reference template.

In others words, the problem is to find the incremental transformation of $\mathbf{z}$, $\boldsymbol{\xi}$ and $\boldsymbol{\gamma}$, that minimizes the sum of squared differences over all the pixels of the cost function

$$c = \min_{\widehat{\mathbf{x}} \circ \mathbf{x}(\mathbf{z}, \boldsymbol{\xi}, \boldsymbol{\gamma})} \frac{1}{2} \sum_{i=1}^{\mathcal{R}} (f_i(\widehat{\mathbf{x}} \circ \mathbf{x}(\mathbf{z}, \boldsymbol{\xi}, \boldsymbol{\gamma}), \mathbf{p}_i))^2 = \frac{1}{2} \|\mathbf{f}_i(\widehat{\mathbf{x}}\mathbf{x}(\mathbf{z}, \boldsymbol{\xi}, \boldsymbol{\gamma}))\|^2 \tag{3.3}$$

where

$$f_i(\widehat{\mathbf{x}} \circ \mathbf{x}(\mathbf{z}, \boldsymbol{\xi}, \boldsymbol{\gamma})) = I'(\mathbf{w}(\widehat{\mathbf{H}}\mathbf{H}(\mathbf{z}), \widehat{\xi} + \xi, \widehat{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{p}_i)) - I(\mathbf{p}_i) \tag{3.4}$$

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

$$\begin{array}{rcl} \widehat{\mathbf{H}} & \leftarrow & \widehat{\mathbf{H}}\mathbf{H}(\mathbf{z}) \\ \widehat{\xi} & \leftarrow & \widehat{\xi} + \xi \\ \widehat{\boldsymbol{\gamma}} & \leftarrow & \widehat{\boldsymbol{\gamma}} + \boldsymbol{\gamma} \end{array} \tag{3.5}$$

Like in [MBMR06], the incremental homography and the intrinsic camera parameters updated are paremetrized with local coordinates of the Lie algebra $\sim\!\!\lessdot(3)$.

Note that the set of parameters does not necessarily correspond to the true parameters. However, this is not important since our goal is to align the images to obtain a visual tracking of the plane.

Also note that, due to image discretisation, we will have to calculate an approximate intensity in the new position. Several standard techniques exist such as nearest neighbour, bilinear interpolation, cubic interpolation. Even though nearest neighbour interpolation has the advantage of being fast and cubic interpolation is a higher order approximation, we observed that bilinear interpolation gave much better results in our tracking tests, for this reason, in all the following experiments, we will use bilinear interpolation.

3.2.2.1 Application of the Efficient Second-order Method (ESM)

The aim is to minimise the objective criterion defined previously in (3.3) in an accurate and robust manner. As this is a non-linear function of the unknown parameters, an iterative procedure is employed. Let $\mathbf{x} = (\mathbf{z}, \boldsymbol{\xi}, \boldsymbol{\gamma})$ be the state vector and let $\widetilde{\mathbf{x}} = (\widetilde{\mathbf{z}}, \widetilde{\boldsymbol{\xi}}, \widetilde{\boldsymbol{\gamma}})$ be the optimal increment such that we obtain the exact solution of the problem: $\mathbf{x} = \widehat{\mathbf{x}}^{-1}\overline{\mathbf{x}} = \tilde{\mathbf{x}}$. Therefore, the objective function is minimized by: $\nabla_{\mathbf{x}}\mathbf{f}_i(\mathbf{x})|_{\mathbf{x}=\widetilde{\mathbf{x}}} = \mathbf{0}$, where $\nabla_{\mathbf{x}}$ is the gradient operator with respect to the unknown parameters and there exists a stationary point $\mathbf{x} = \widetilde{\mathbf{x}}$ which is the global minimum of the cost function.

Since both the reference image and current image are available it is possible to use the efficient second-order approximation method (ESM) [Mal04] to solve the optimization problem. In this case the current and reference Jacobians are

$$\mathbf{J}(\mathbf{0}) = \mathbf{J}_{\mathbf{I}'}\mathbf{J}_{\mathbf{w}}\left[\begin{array}{ccc} \mathbf{J}_{\mathbf{c}}\mathbf{J}_{\mathbf{s_H}}\mathbf{J}_{\mathbf{H_z}}(\mathbf{0}) & \mathbf{J}_{\mathbf{c}_\xi}(\mathbf{0}) + \mathbf{J}_{\mathbf{c}}\mathbf{J}_{\mathbf{s}}\mathbf{J}_{\mathbf{c}_\xi^{-1}}(\mathbf{0}) & \mathbf{J}_{\mathbf{c}_\gamma}(\mathbf{0}) + \mathbf{J}_{\mathbf{c}}\mathbf{J}_{\mathbf{s}}\mathbf{J}_{\mathbf{c}_\gamma^{-1}}(\mathbf{0}) \end{array}\right] \tag{3.6}$$

$$\mathbf{J}(\widetilde{\mathbf{x}}) = \mathbf{J}_{I}\mathbf{J}_{\widetilde{\mathbf{w}}}\left[\begin{array}{ccc} \mathbf{J}_{\widetilde{\mathbf{c}}}\mathbf{J}_{\widetilde{\mathbf{s}}_\mathbf{H}}\mathbf{J}_{\widetilde{\mathbf{H}}_\mathbf{z}}(\widetilde{\mathbf{x}}) & \mathbf{J}_{\widetilde{\mathbf{c}}_\xi}(\widetilde{\mathbf{x}}) + \mathbf{J}_{\widetilde{\mathbf{c}}}\mathbf{J}_{\widetilde{\mathbf{s}}}\mathbf{J}_{\widetilde{\mathbf{c}}_\xi^{-1}}(\widetilde{\mathbf{x}}) & \mathbf{J}_{\widetilde{\mathbf{c}}_\gamma}(\widetilde{\mathbf{x}}) + \mathbf{J}_{\widetilde{\mathbf{c}}}\mathbf{J}_{\widetilde{\mathbf{s}}}\mathbf{J}_{\widetilde{\mathbf{c}}_\gamma^{-1}}(\widetilde{\mathbf{x}}) \end{array}\right] \tag{3.7}$$

Thanks to the left invariance property $\mathbf{J}(\widetilde{\mathbf{x}})\widetilde{\mathbf{x}} \approx \mathbf{J}(\mathbf{0})\widetilde{\mathbf{x}}$, the reference jacobian $\mathbf{J}(\widetilde{\mathbf{x}})\widetilde{\mathbf{x}}$ can be partially calculated without knowing the true value of $\widetilde{\mathbf{x}}$ (this is the basis of the ESM algorithm). Then, our second-order least-squares minimiser is the solution to

$$\widetilde{\mathbf{x}} = \left(\left(\frac{\mathbf{J}_I + \mathbf{J}_{I'}}{2}\right)\mathbf{J}_{\mathbf{w}}\left[\begin{array}{ccc} \mathbf{J}_{\mathbf{c}}\mathbf{J}_{\mathbf{s_H}}\mathbf{J}_{\mathbf{H_z}}(\mathbf{0}) & \mathbf{J}_{\mathbf{c}_\xi}(\mathbf{0}) + \mathbf{J}_{\mathbf{c}}\mathbf{J}_{\mathbf{s}}\mathbf{J}_{\mathbf{c}_\xi^{-1}}(\mathbf{0}) & \mathbf{J}_{\mathbf{c}_\gamma}(\mathbf{0}) + \mathbf{J}_{\mathbf{c}}\mathbf{J}_{\mathbf{s}}\mathbf{J}_{\mathbf{c}_\gamma^{-1}}(\mathbf{0}) \end{array}\right]\right)^{+}\mathbf{f}(\mathbf{0}) \tag{3.8}$$

The symbol '+' indicates the pseudo-inverse matrix. $\mathbf{J}_I$ is the Jacobian of the image reference and therefore only needs to be calculated once. The rest of Jacobians are recomputed at each iteration.

The current Jacobian and reference Jacobian are detailed in the Appendix A.

3.3 Experimental Results

In order to assess the performance of the proposed method, we have performed experiments with synthetic and real data. For the synthetic data we have created image sequences considering different configurations of perspective camera and mirror shape (parabolic, hyperbollic and elliptic). For the real data we have tested the algorithm with the two kinds of catadioptric omnidirectional cameras (parabolic and hyperbolic) that we have at INRIA Sophia Antipolis Méditerranée (See Figure 3.6).

Also, the proposed method has been compared with the closest work in the literature, that of C. Mei *et al.* [MBMR08] (calibrated visual tracking with an omnidirectional camera). This comparison serves to highlight the differences between the proposed method and [MBMR08], where since a calibrated camera is considered the intrinsic parameters are known. In this comparison we see the increased complexity of our problem when we use the approach in [MBMR08] with an uncalibrated camera where the intrinsic parameters are unknown. Our tracking approach works

both when a calibrated camera or an uncalibrated camera is considered. However, the approach in [MBMR08] fails when an uncalibrated camera is considered. This highlights the difficulties that arise when dealing with an uncalibrated camera.

Figure 3.6: *Left. A para-catadioptric camera consisting of a S80 parabolic mirror from RemoteReality with telecentric lens and a perspective camera with an image resolution of 1024 × 768. Right. A hyper-catadioptric camera consisting of a hyperbolic mirror from Accowle Vision and a perspective camera with an image resolution of 1024 × 768.*

3.3.1 Synthetic data

In the following we will show a series of experiments with synthetic data. We have simulated the cases of an omnidirectional camera using a parabolic, hyperbolic and elliptical mirror. The planar mirror was not considered because it does not make sense to build an omnidirectional camera using a planar mirror. However, this case can be approximated to the perspective case wich has been vastly used for tracking tasks. Some works with uncalibrated tracking using a perspective camera can be found in http://esm.gforge.inria.fr/ESM.html.

3.3.1.1 Omnidirectional Parabolic camera

Our experimental setup consists of an image sequence composed of 40 images. To create this sequence we transformed a real parabolic image (See Figure 3.7) assuming constant intrinsic parameters such as a catadioptric camera with a parabolic

mirror ($\xi = 1$), a generalized focal length $\beta_1 = -250$, $\beta_2 = -250$ and an image center $u_0 = 512$ and $v_0 = 384$. The images were spaced by the homography parameters depicted in Figure 3.8. There is noise due to discretisation. An approximate intensity in the new position is computed by bilinear interpolation.

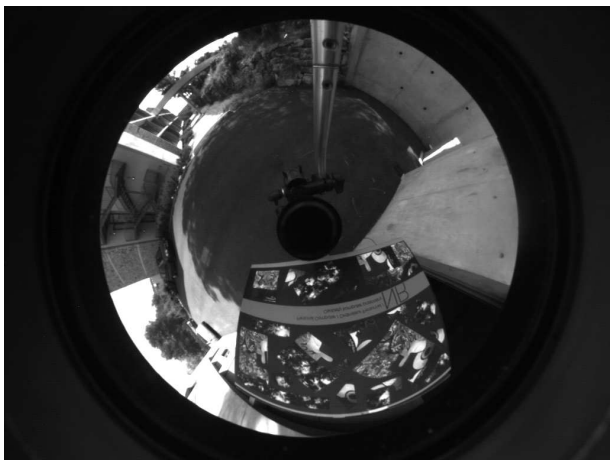


Figure 3.7: *Catadioptric image. Image taken with a para-catadioptric camera consisting of a parabolic mirror, telecentric lens and a perspective camera.*

In order to compare our algorithm with the algorithm proposed by C. Mei *et al.* [MBMR08], we have performed two experiments.

First experiment

In this experiment, we have considered known intrinsic parameters. To start the minimization the intrinsic parameters for both methods were $\xi = 1$, $\beta_1 = -250$, $\beta_2 = -250$, $u_0 = 512$ and $v_0 = 384$. For the method in [MBMR08] these parameters are constants. It only computes the homography parameters. The initial guess for the homography parameters was given by the 3×3 identity matrix. Figure 3.9 shows the reprojection error (RMS) using the proposed visual tracking algorithm in [MBMR08] (left) and for the proposed method in this chapter (right). For the entire sequence, the reprojection errors are indistinguishable. Figure 3.10 shows some images of the test sequence with the tracked image region marked in red.

Second experiment

This experiment tests the goal of our algorithm, therefore, we have considered unknown the intrinsic parameters to start the minimization method, so we gave an initial guess of $\xi = 0.75$, $\beta_1 = -350$, $\beta_2 = -340$, $u_0 = 497$ and $v_0 = 402$. The initial guess for the homography parameters was given by the 3×3 identity matrix. Figure

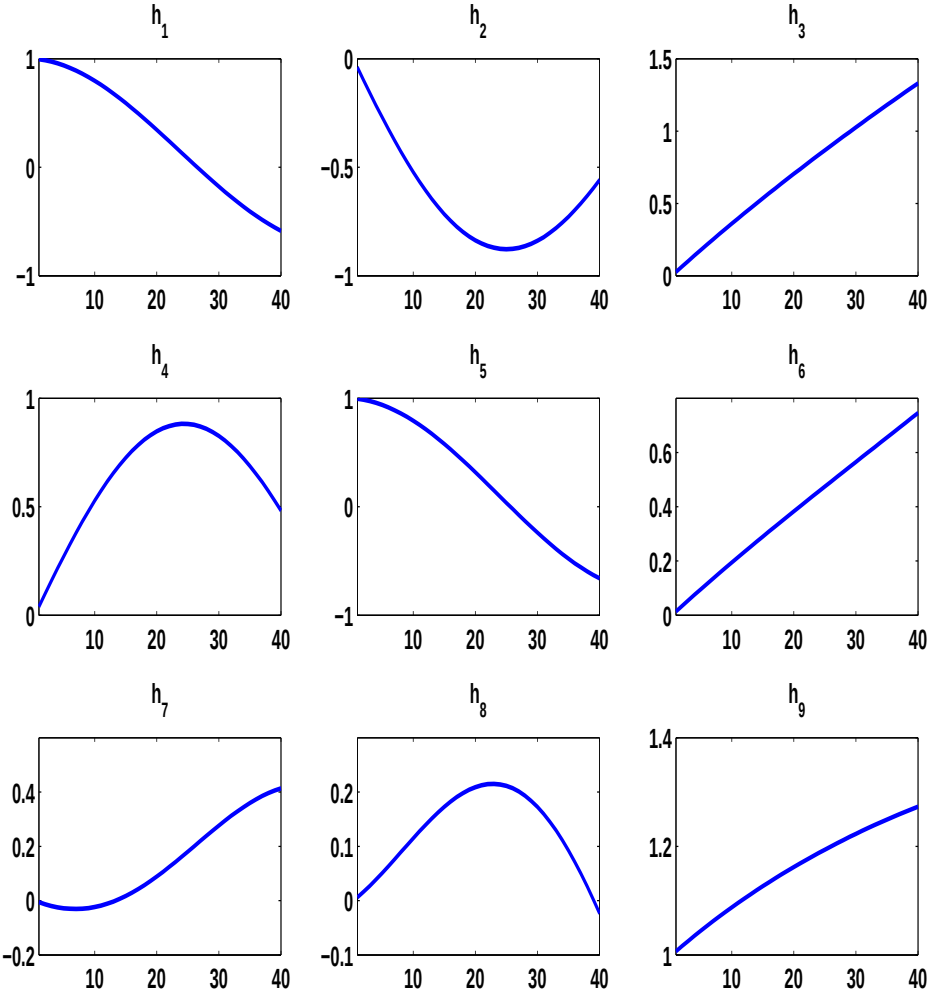


Figure 3.8: *Homography parameters. These parameters were used to create the para-catadioptric image sequence of the first experiment with synthetic data in Section 3.3.1.1.*

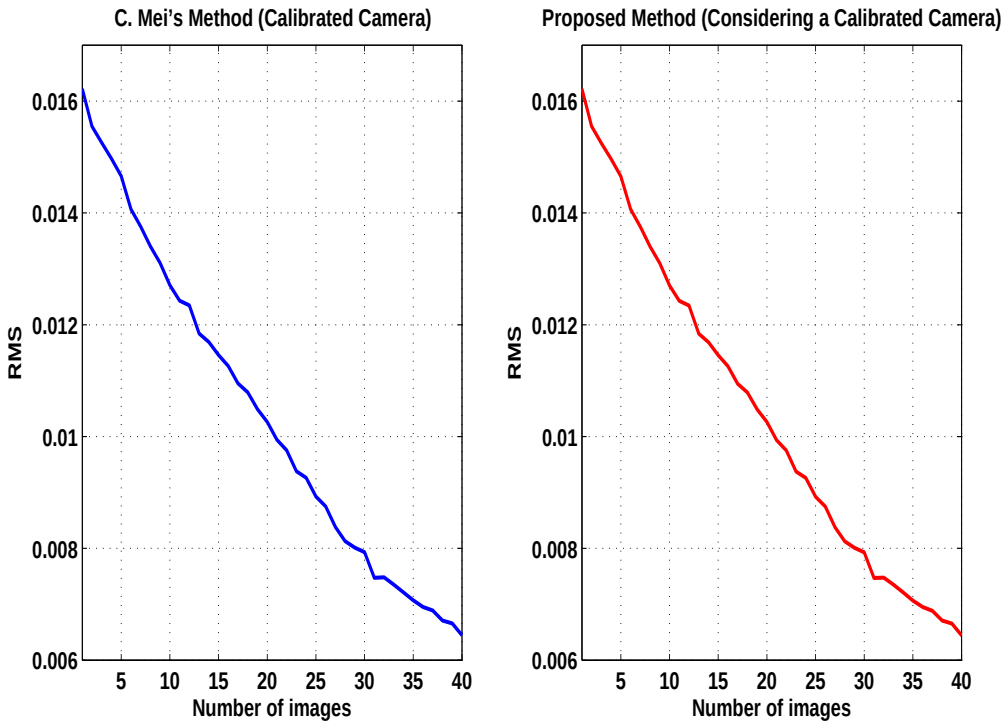


Figure 3.9: *Reprojection error. This figure shows the result of tracking 40 para-catadioptric images in the first experiment with synthetic data in Section 3.3.1.1.*

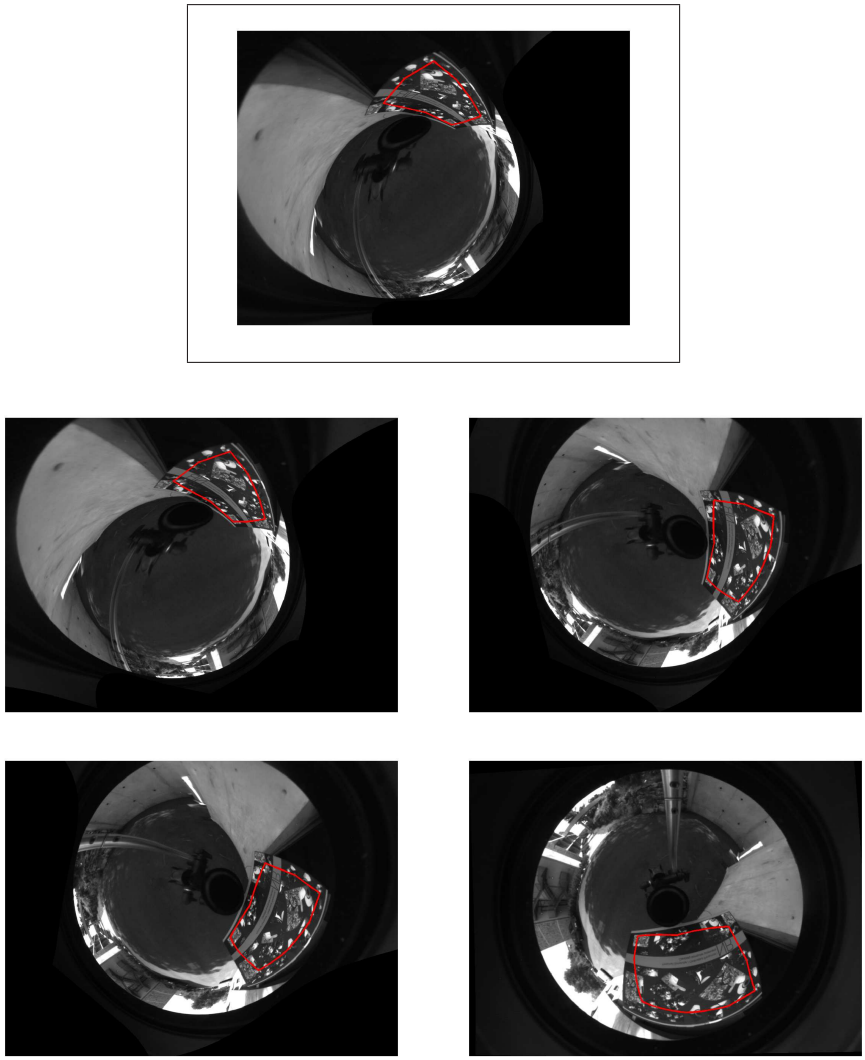


Figure 3.10: *Visual tracking of synthetic para-catadioptric images in the first experiment of Section 3.3.1.1. Top. Reference plane. Bottom. Some images from the tracking output sequence using our method. The same output is obtained using the method in [MBMR08].*

3.11 shows the reprojection error (RMS) using the visual tracking algorithm proposed in [MBMR08] (left) and for the method proposed in this paper (right). The reprojection erros are very different. Figure 3.12 shows some images of the test sequence with the tracked image region marked in blue for the method in [MBMR08]. Figure 3.13 shows some images of the test sequence with the tracked image region marked in red for our method. As we can see, the method proposed in [MBMR08] is not able to track accurately the expected path along the sequence. On the other hand, even if the camera is not calibrated, the proposed algorithm is capable of tracking accurately the chosen plane along the sequence.

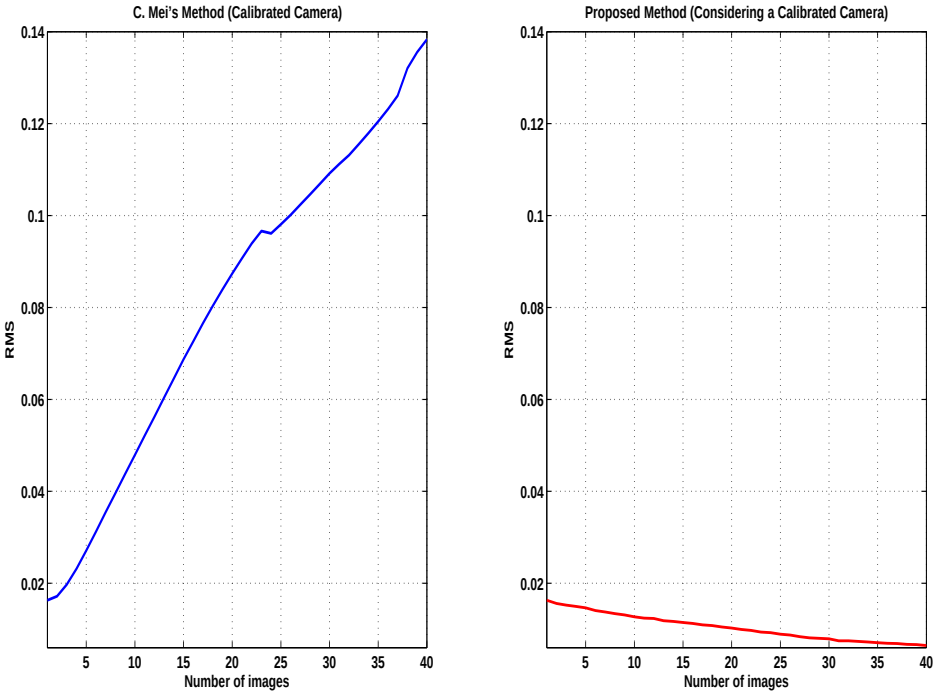


Figure 3.11: *Reprojection error. This figure shows the result of tracking 40 images in the second experiment with synthetic data in Section 3.3.1.1.*

3.3.1.2 Omnidirectional Hyperbolic or Elliptical camera

In this case, the experimental setup consists of a image sequence composed of 120 images. To create this sequence we transformed the image from Figure 3.7 assuming constant intrinsic parameters such as a hyperbolic or elliptical camera ($\xi = 0.89$), a generalized focal length $\beta_1 = -400$, $\beta_2 = -370$ and an image center $u_0 = 522$ and $v_0 = 369$. The images were spaced by the homography parameters depicted in Figure 3.14. The choice of $\xi = 0.89$ was made because from our experience of omnidirectional camera calibration the ranges of ξ for these divices are between 0.80 to 0.95. Again, there is noise due to the approximation of intensity by bilinear interpolation.

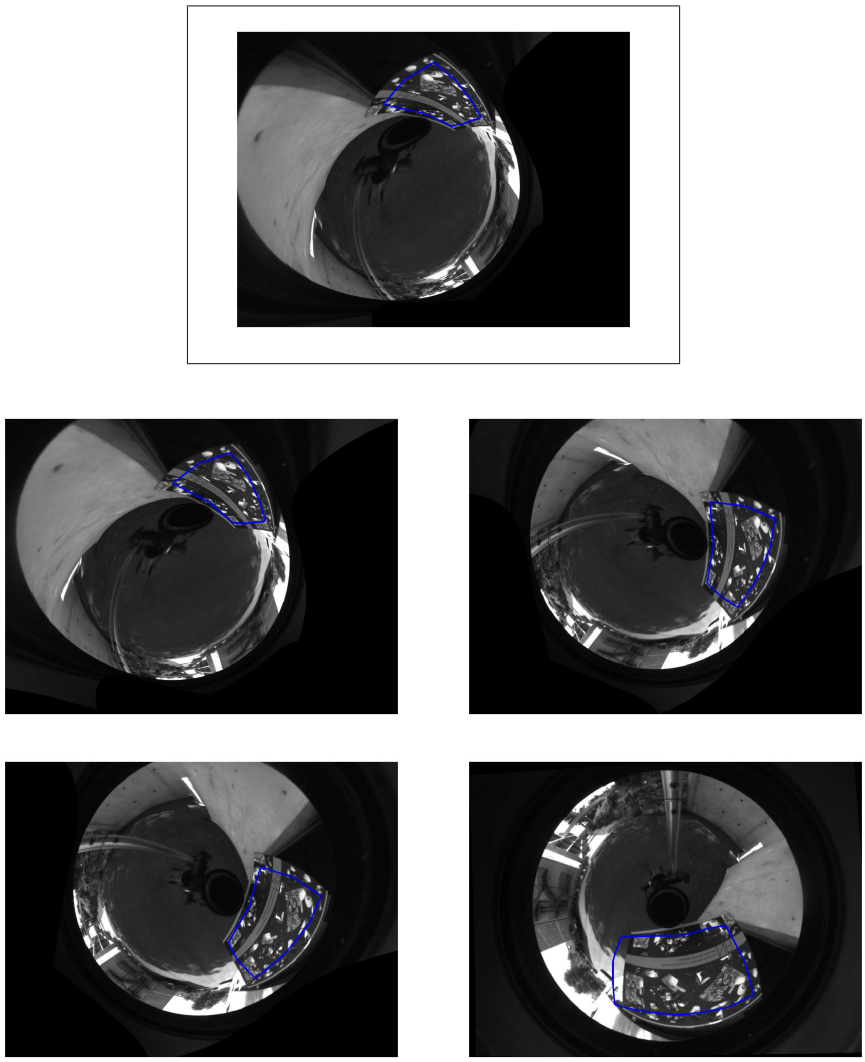


Figure 3.12: *Visual tracking of synthetic para-catadioptric images in the second experiment of Section 3.3.1.1. Top. Reference plane. Bottom. Some images from the tracking output sequence using the method proposed in [MBMR08].*

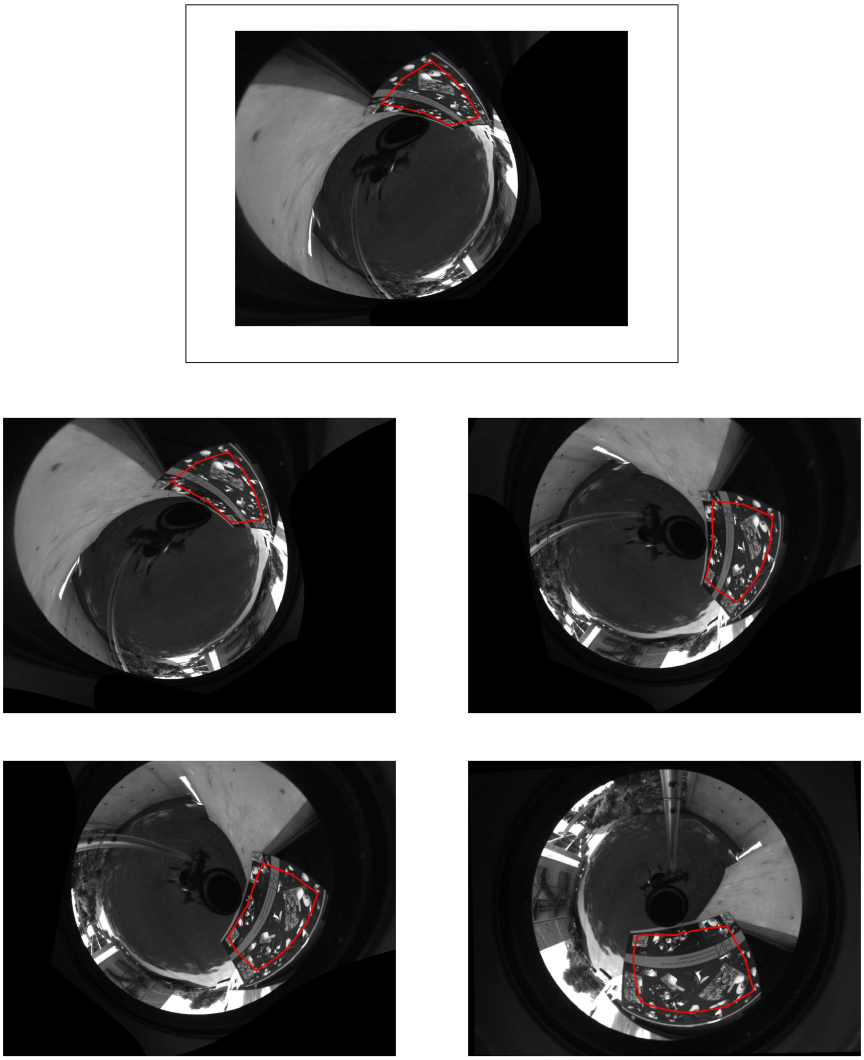


Figure 3.13: *Visual tracking of synthetic para-catadioptric images in the second experiment of Section 3.3.1.1. Top. Reference plane. Bottom. Some images from the tracking output sequence using the proposed visual tracking algorithm.*

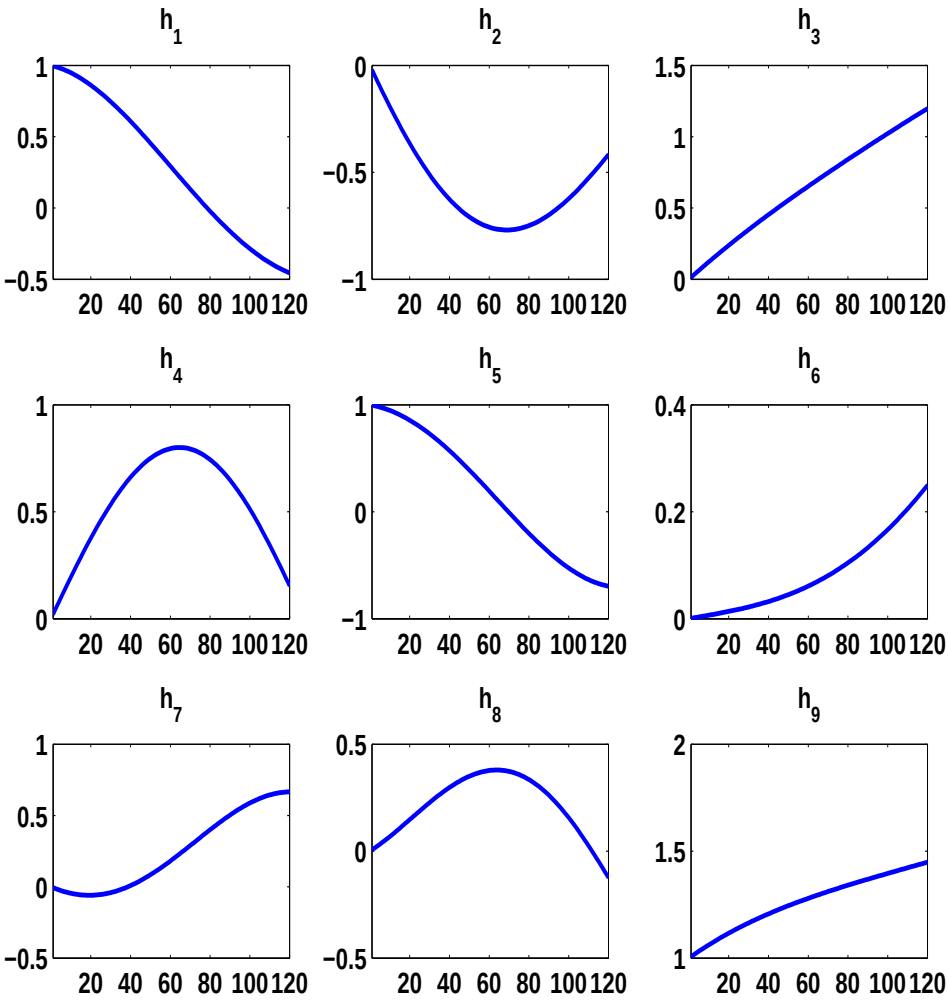


Figure 3.14: *Homography parameters. These parameters were used to create the hyper(elliptic)-catadioptric image sequence of the first experiment with synthetic data in Section 3.3.1.2.*

Like in the previous section, we conducted two experiments in order to compare our algorithm with the algorithm proposed by C. Mei *et al.* [MBMR08].

First experiment

In this experiment, we have considered known intrinsic parameters. To start the

minimization the intrinsic parameters for both methods were $\xi = 0.89$, $\beta_1 = -400$, $\beta_2 = -370$, $u_0 = 522$ and $v_0 = 369$. For the method in [MBMR08] these parameters are constants. It only computes the homography parameters. The initial guess for the homography parameters was given by the 3×3 identity matrix. Figure 3.15 shows the reprojection error (RMS) using the proposed visual tracking algorithm in [MBMR08] (left) and for the proposed method in this chapter (right). For the entire sequence, the reprojection errors are indistinguishable. Figure 3.16 shows some images of the test sequence with the tracked image region marked in red.

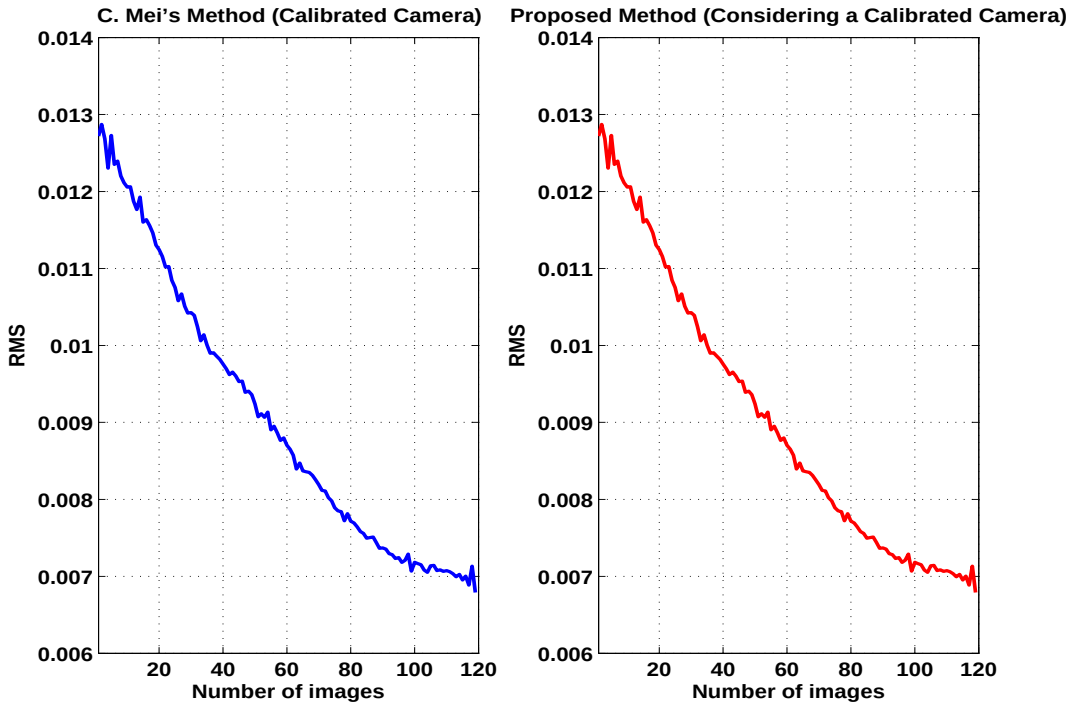


Figure 3.15: *Reprojection error. This figure shows the result of tracking 120 images in the first experiment with synthetic data in Section 3.3.1.2.*

Second experiment

For this experiment we considered the case with unknown intrinsic parameters. To start the minimization method we gave an initial guess of $\xi = 0.5$, $\beta_1 = -100$, $\beta_2 = -100$, $u_0 = 500$ and $v_0 = 390$. The initial guess for the homography parameters was given by the 3×3 identity matrix. Figure 3.17 shows the reprojection error (RMS) using the proposed visual tracking algorithm in [MBMR08] (left) and for the proposed method in this chapter (right). The reprojection erros are very different. Figure 3.18 shows some images of the test sequence with the tracked image region marked in blue for the method in [MBMR08]. Figure 3.19 shows some images of the test sequence with the tracked image region marked in red for the proposed method. Again, as the reader can see, the method proposed in [MBMR08] is not able to track

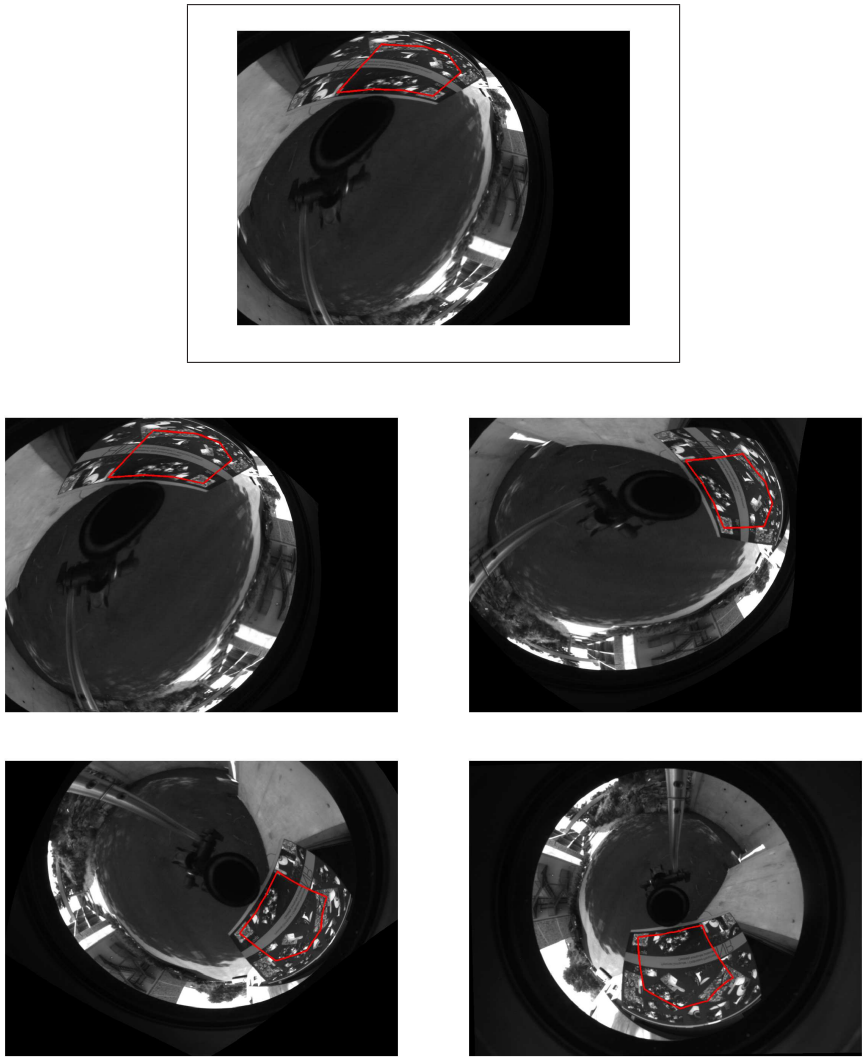


Figure 3.16: *Visual tracking of synthetic hyper(elliptic)-catadioptric images in the first experiment of Section 3.3.1.2. Top. Reference plane. Bottom. Some images from the tracking output sequence using our method. The same output is obtained using the method in [MBMR08].*

accurately the expected path along the sequence. On the other hand, even if the camera is not calibrated, the proposed algorithm is capable of tracking accurately the chosen plane along the sequence.

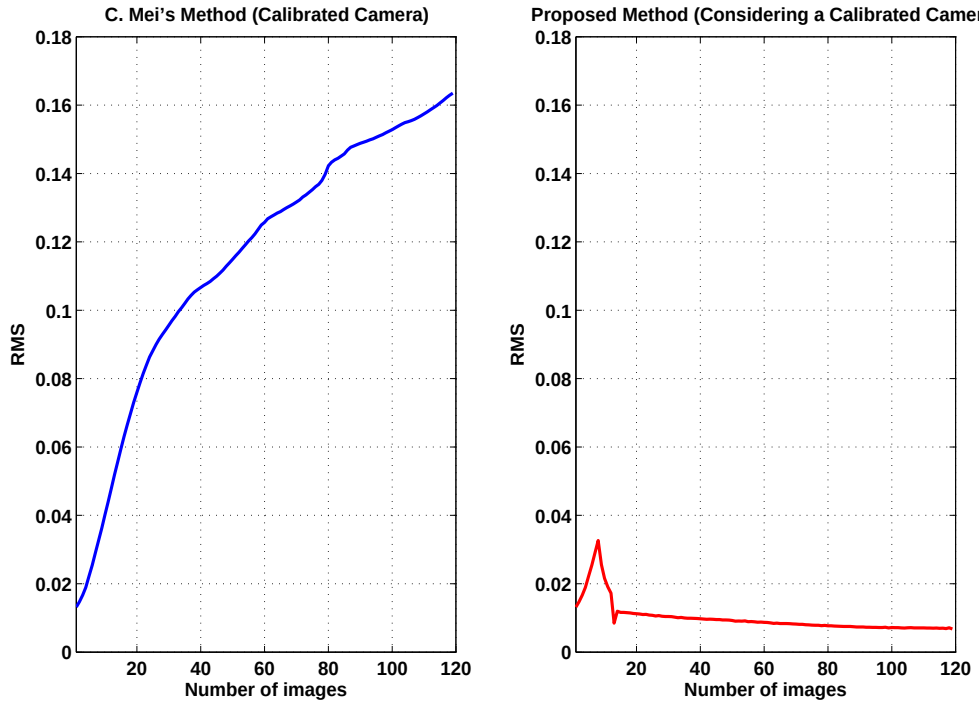


Figure 3.17: *Reprojection error. This figure shows the result of tracking 120 images in the second experiment with synthetic data in Section 3.3.1.2.*

3.3.2 Real data

In the following we will show a series of experiments with real data. Two types of omnidirectional cameras have been considered (parabolic and hyperbolic).

3.3.2.1 Omnidirectional Parabolic camera

Our experimental setup consists of two image sequences (120 and 1000 images each) of size 1024×768 combining a perspective camera with a telecentric lens and a parabolic mirror (See left side of Figure 3.6). Here, our algorithm has been compared against the algorithm in [MBMR08] just for the first sequence. Also, we considered only the case with unknown parameters (uncalibrated camera).

First sequence

To start the minimization method we gave an initial guess of $\xi = 0.9$, $\beta_1 = -80$,

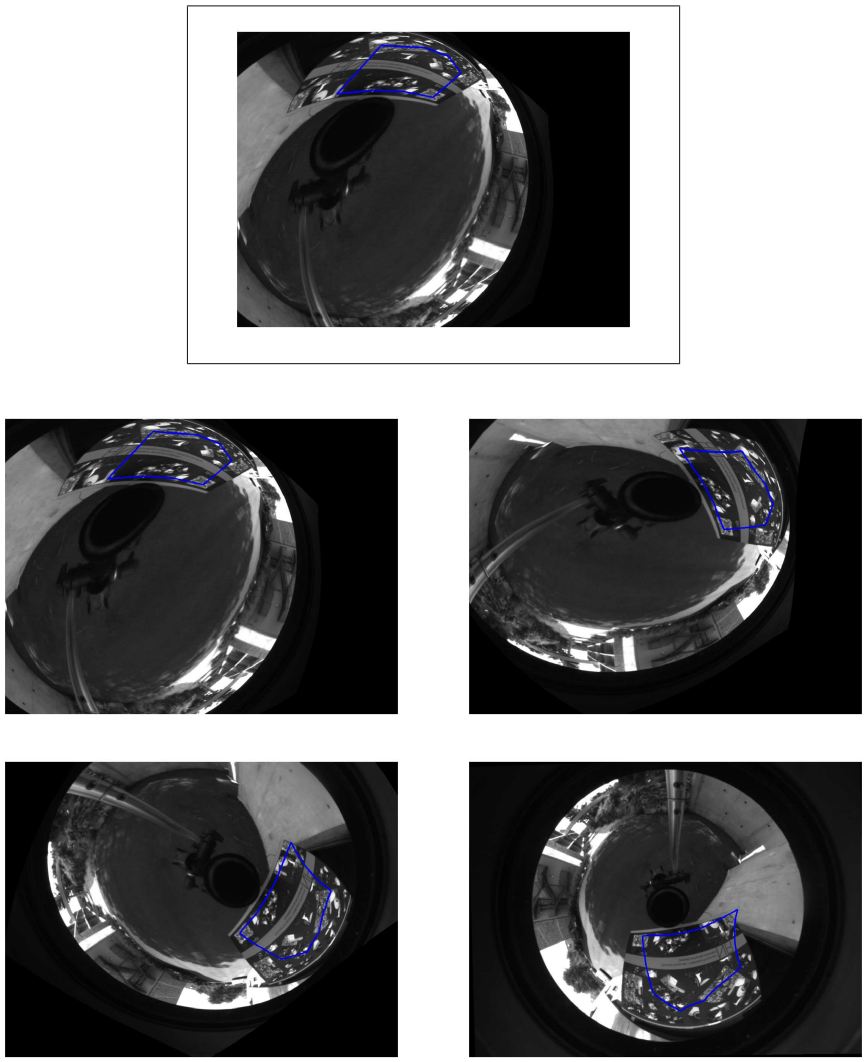


Figure 3.18: *Visual tracking of synthetic hyper(elliptic)-catadioptric images in the first experiment of Section 3.3.1.2. Top. Reference plane. Bottom. Plane tracked using the method proposed in [MBMR08].*

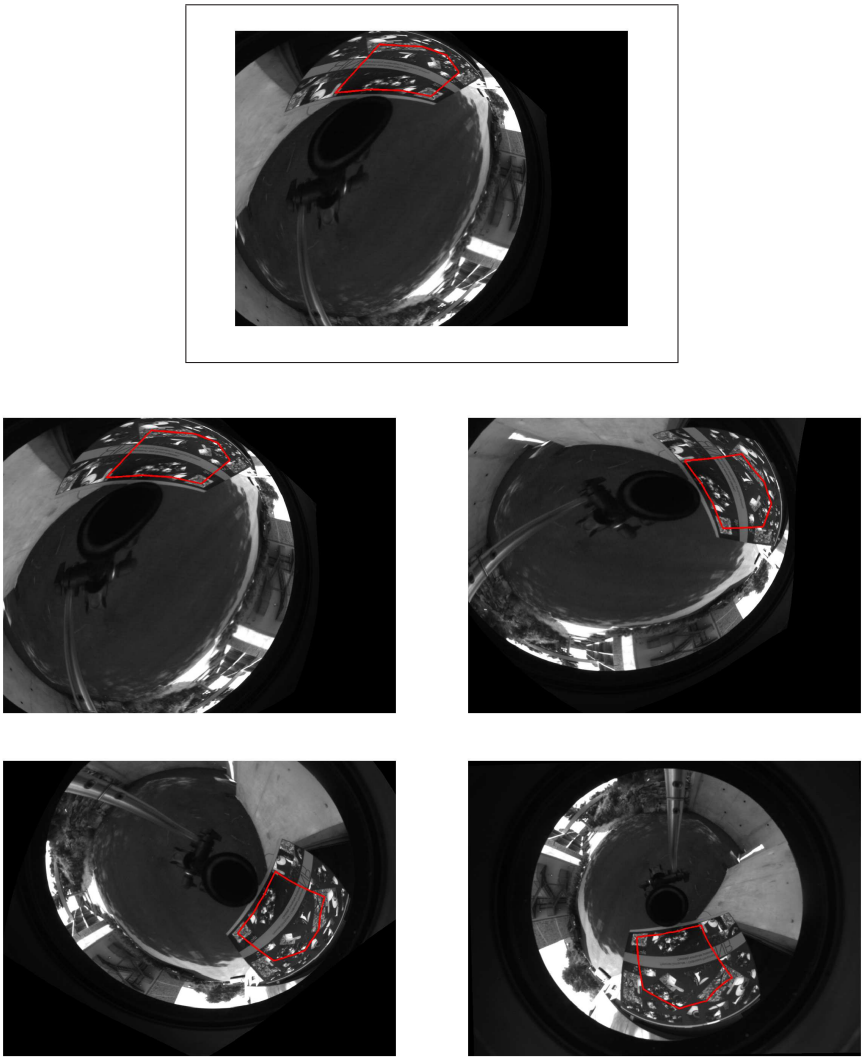


Figure 3.19: *Visual tracking of synthetic hyper(elliptic)-catadioptric images in the first experiment of Section 3.3.1.2. Top. Reference plane. Bottom. Some images from the tracking output sequence using the proposed visual tracking algorithm.*

$\beta_2 = -80$, $u_0 = 512$ and $v_0 = 384$. The initial guess for the homography parameters was given by the 3×3 identity matrix. Figure 3.20 shows the reprojection error (RMS) using the proposed visual tracking algorithm in [MBMR08] (left) and for the proposed method in this chapter (right). Figure 3.21 shows some images of the real sequence with the tracked image region marked in blue for the method in [MBMR08]. Figure 3.22 shows some images of the real sequence with the tracked image region marked in red for our algorithm. Again, like the synthetic data, the proposed method in [MBMR08] is not robust for tracking the sequence with unknown intrinsic parameters. That means that the homography matrix is not enough to minimise the reprojection error between the reference image and the current image while the displacement is increasing. On other hand, even if the camera is not calibrated, the proposed algorithm is capable of tracking accurately the choosen plane along the sequence.

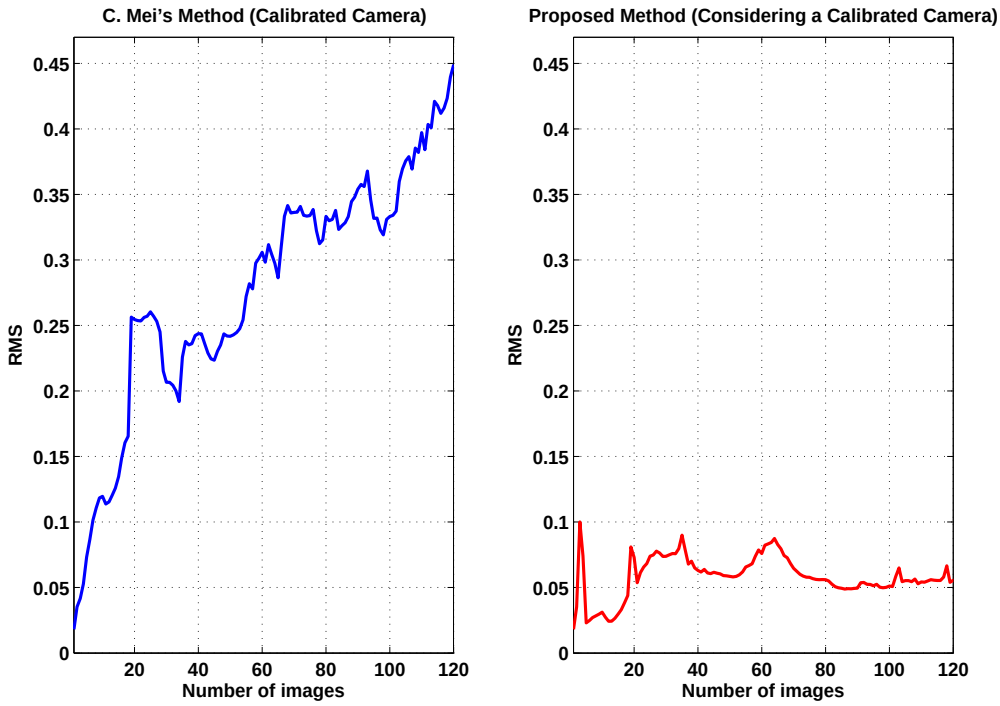


Figure 3.20: *Reprojection error. This figure shows the result of tracking 120 images in the first experiment with real data in Section 3.3.2.1.*

Second sequence

To start the minimization method we gave an initial guess of $\xi = 1.0$, $\beta_1 = -100$, $\beta_2 = -100$, $u_0 = 512$ and $v_0 = 384$. The initial guess for the homography parameters was given by the identity 3×3 matrix. Figure 3.23 shows the reprojection error (RMS) for the proposed method in this chapter. Figure 3.24 shows some images of the real sequence with the tracked image region marked in red for our algorithm.

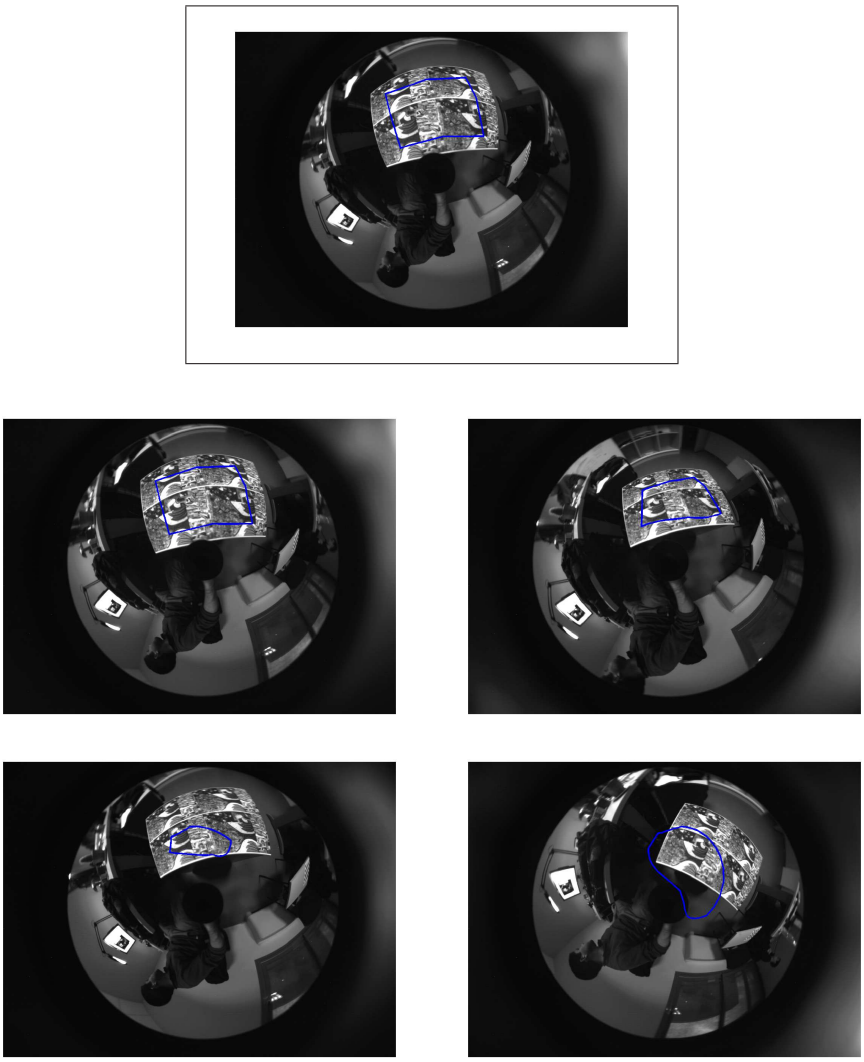


Figure 3.21: *Visual tracking of real para-catadioptric images in the first experiment of Section 3.3.2.1. Top. Reference plane. Bottom. Some images from the tracking output sequence using the method proposed in [MBMR08].*

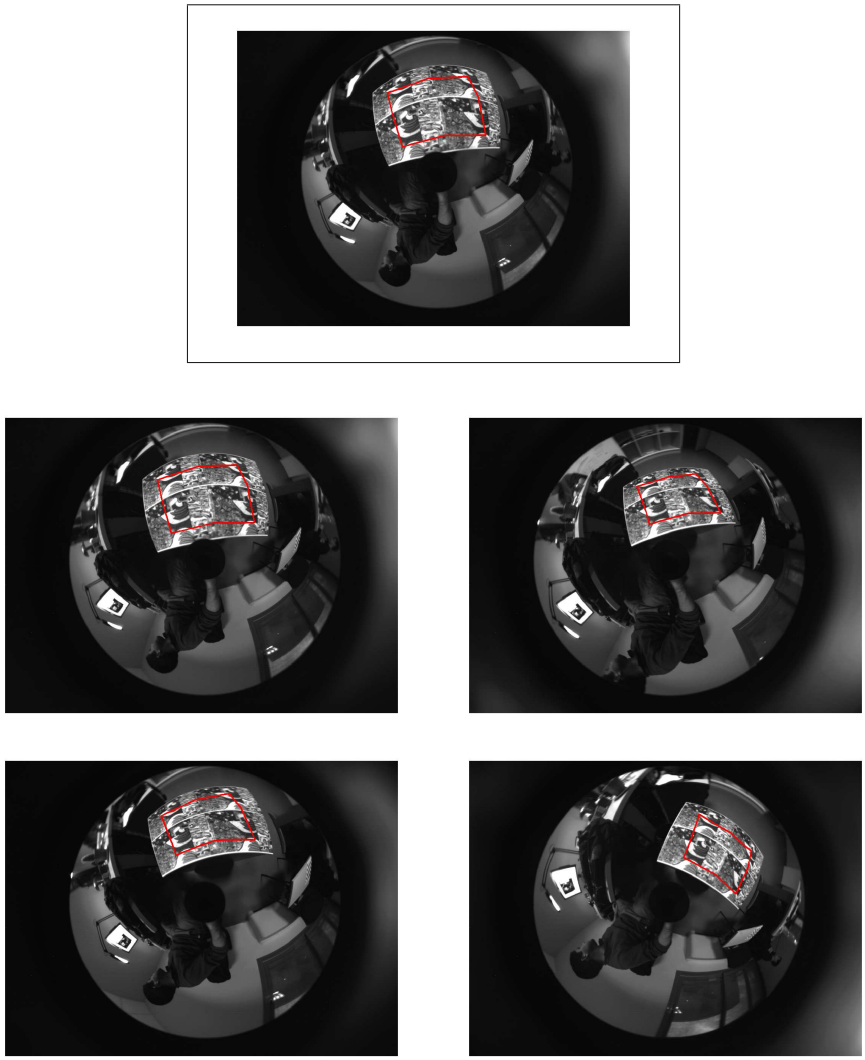


Figure 3.22: *Visual tracking of real para-catadioptric images in the first experiment of Section 3.3.2.1. Top. Reference plane. Bottom. Some images from the tracking output sequence using the proposed visual tracking algorithm.*

Again, the proposed algorithm is capable of tracking accurately the choosen plane along the sequence.

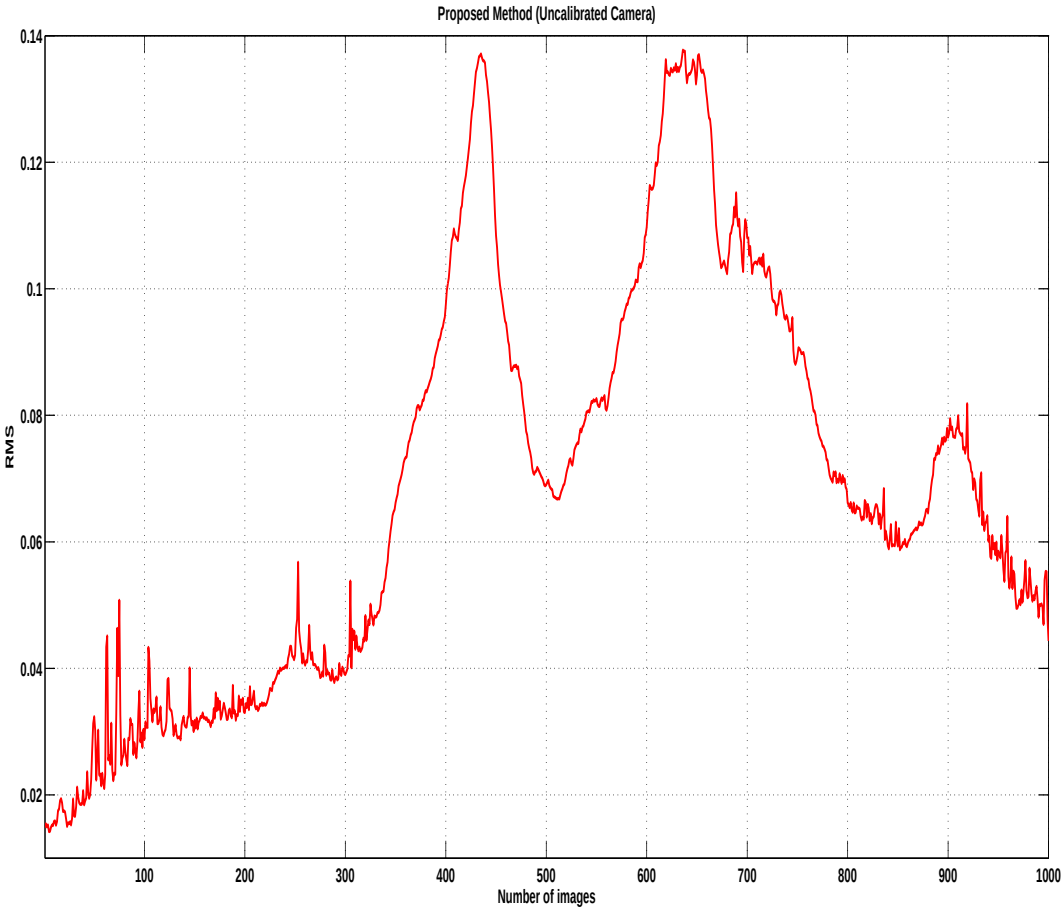


Figure 3.23: *Reprojection error. This figure shows the result of tracking 120 images in the second experiment with real data in Section 3.3.2.1.*

3.3.2.2 Omnidirectional Hyperbolic Camera

Our experimental setup consists of one image sequence with 300 images of size 640 $\times$ 480 combining a perspective camera with a hyperbolic mirror (See Figure 3.6). In this experiment, we have only considered the case with unknown parameters (uncalibrated camera).

The minimization method was initialised with the following guesses for intrinsic parameters $\xi = 0.9$, $\beta_1 = -100$, $\beta_2 = -100$, $u_0 = 320$ and $v_0 = 240$. The initial guess for the homography parameters was given by the 3×3 identity matrix. Figure 3.25 shows the reprojection error (RMS) using our algorithm. Figure 3.26 shows some images of the real sequence with the tracked image region marked in red for our algorithm. Again, like synthetic data, the proposed algorithm is capable of tracking accurately the choosen plane along the sequence.

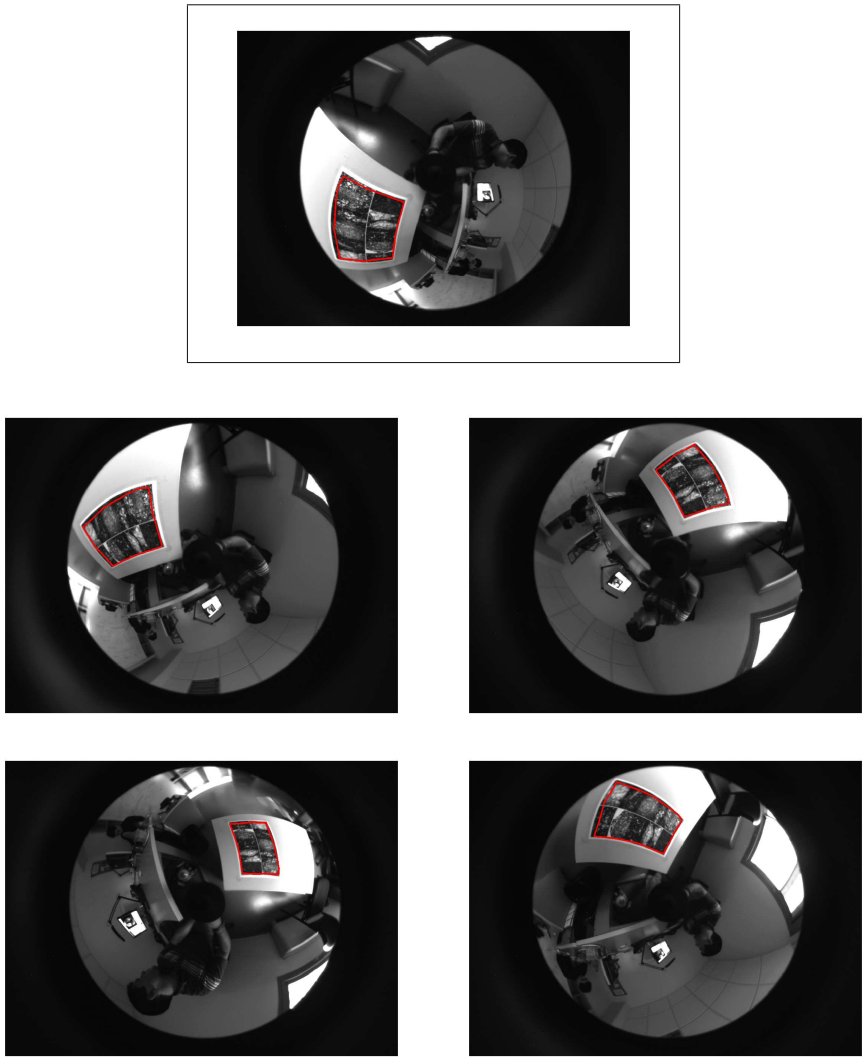


Figure 3.24: *Visual tracking of real para-catadioptric images in the second experiment of Section 3.3.2.1. Top. Reference plane. Bottom. Some images from the tracking output sequence using the proposed visual tracking algorithm.*

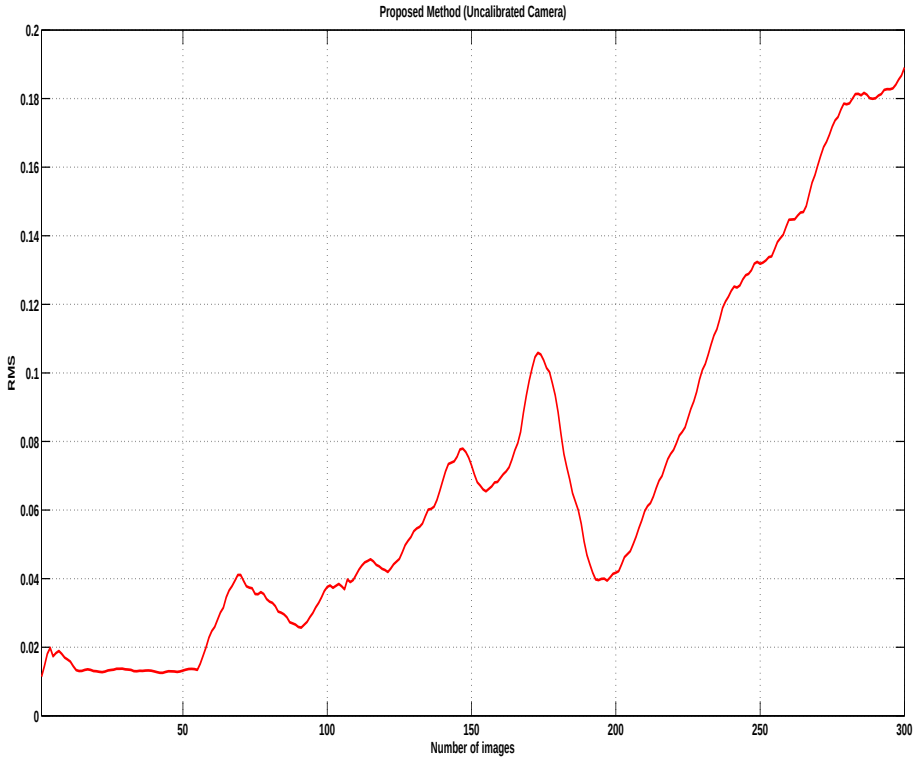


Figure 3.25: *Reprojection error. This figure shows the result of tracking 300 images in the experiment with real data in Section 3.3.2.2. The error is increasing because the light changes in the images (See Figure 3.26).*

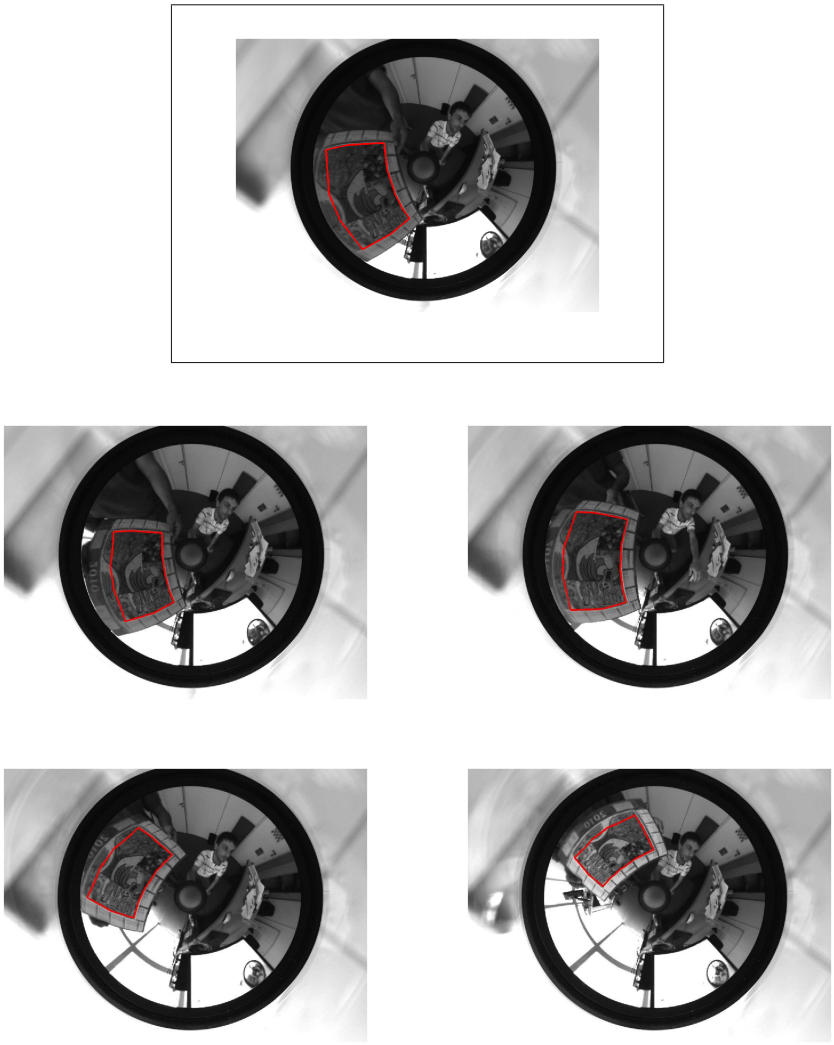


Figure 3.26: *Visual tracking of real hyper-catadioptric images in the experiment of Section 3.3.2.2. Top. Reference plane. Bottom. Some images from the tracking output sequence using the proposed visual tracking algorithm.*

3.4 Conclusion

In this chapter, we have shown how to efficiently track a plane in an omnidirectional image without requiring the prior calibration of the sensor. A set of required parameters are estimated on-line for each new image to align the current image with a reference template. The approach is novel because the estimated parameters are integrated into a single global warping function and we developed the Jacobian matrix of this warping function in easy modular parts. Furthermore, the efficient second order minimisation technique was applied, which allowed us to minimise a highly redundant non-linear function with high precision.

In the following chapter, we will detail how to self-calibrate omndirectional cameras by using several of the tracked views. This also enables us to fix the values being estimated, providing a faster and more robust algorithm. It will be shown that the unified projection model leads to a flexible calibration approach and is well adapted to central catadioptric calibration for precise robotic applications.

Chapter 4

Direct On-line Self-Calibration

In this chapter we will present a calibration method that can be carried out using the same images required for performing the visual task. The algorithm works directly with the image intensity. It stays valid for all central catadioptric systems and needs no prior knowledge about extrinsic, camera, lens or mirror parameters. The only requirement is to select a plane in the image scene. Then, on-line self-calibration is performed by tracking the target region (selected plane) in several incoming images.

Also, we will formalize the uniqueness of the solution for the calibration problem of central catadioptric omnidirectional cameras. The calibration problem is formalized by using a unified model that is valid for all central catadioptric omnidirectional cameras and can be extended to traditional cameras when a planar mirror is considered. Besides, we show that the uniqueness of the problem can be derived from the solution of non-linear equations.

4.1 Self-calibration of Omnidirectional cameras

Visual sensor calibration is the problem of determining the parameters of the transformation between the 3D information of the imaged object in space and the 2D observed image. Such a relationship is mandatory for 3D vision (3D information can be infered from 2D information, and vice-versa). More precisely, in order to obtain meaningful geometric information from a camera, two calibration procedures must be completed. The first is intrinsic calibration, that is, determining the internal camera parameters (e.g., focal length, principal point, and skew coefficients), which affect the image measurements. The second is extrinsic calibration, which is the process of computing the transformation between the camera and a base frame of reference. In a surveillance application, the base frame may be the room or building coordinate system, whereas on a mobile robot, the base frame could be the robot-body frame.

The problem of the self-calibration of omnidirectional cameras has attracted the attention of researchers in the computer vision community because, often, no information about mirror or lens parameters and no calibration objects are available. Then, unlike classical calibration problem (knowledge about the scene), the algorithms for self-calibration make no or few assumptions about the particular structure of the scene being viewed. Instead, they attempt to find the intrinsic parameters and extrinsic parameters of the camera by exploiting constraints imposed over these parameters from epipolar or trilinear relations, from specific camera motions or from correspondences between multiple views. In fact, the word self-calibration for

a visual sensor is applicable when the method allows the system to calibrate (1) automatically, (2) without any semantic knowledge of the observed objects and (3) without a complete knowledge of the self-motion.

Several works have addressed camera self-calibration [BP02, GN98, GD01, Kan00, MP04, FTB02, BV08]. In the following section we briefly review the state of the art of these techniques and set the context for the algorithm that we are proposing for self-calibrating an omnidirectional camera.

4.1.1 Self-calibration techniques

In [Kan00] two self-calibration techniques are proposed for parabolic mirros. The first one is the so-called circle-based self-calibration. The idea is to identify the bounding circle of the omnidirectional image. This can be done manually or automatically by using a predefined threshold, by finding the boundary, and by fitting a circle to the resulting boundary. The center of the circle is deemed to be the image principal point. Since the field of view of the omnidirectional device is known, the camera's parabolic parameter can then be directly computed using the radius of the circle. This technique is simple to implement and does not need any knowledge of the scene or extraction of features in the scene. Only one image is necessary to calibrate the sensor. However, if the circle must be automatically extracted, finding the optimal threshold is difficult due to changing lighting conditions. In addition, a single static threshold may not be sufficient, due to directional lighting that may make one side brighter than the other. The second technique is based on point feature tracks across an omnidirectional image sequence. An example of a collection of feature tracks generated is shown in Figure 4.1. It uses consistency of pairwise correspondence for an imaging device and considers a paraboloid mirror and an orthographic camera. Thus, the method cannot be implemented for different kinds of omnidirectional mirrors. The algorithm works as follows: First, point tracks are obtained using a tracker developed in [ST94a]. Then, it estimates the essential matrix $\mathbf{E}$ from point correspondences for some intrinsic parameters. The calibration parameters are obtained by minimizing the algebraic or geometric error of point correspondences to the epipolar curves using the Nelder-Mead simplex search algorithm [Num02].

In [MP04], the authors presented a para-catadioptric camera self-calibration method from point correspondences and epipolar constraints. The method is philosophically the same as the one in [Kan00] but it provides a closed-form solution to a polynomial eigenvalue problem and therefore no iterative minimizing process has to be used. The closed form solution allows one to apply a 9-point RANSAC robust estimation technique to handle outliers in matches. Even though the method was designed for a parabolic mirror the authors showed that it can be applied (after a linearization of the model) with sufficient accuracy for a hyperbolic mirror. Figure 4.2 shows the outliers detection during the calibration process.

In [GD01], Geyer and Daniilidis performed calibration of a paracatadioptric camera from point correspondences (Figure 4.3). The authors assume that there are three unknown intrinsic parameters: the combined focal length of the mirror and lens and

Figure 4.1: *Kang's method. Example of tracking points in a real office scene for self-calibration porpuses.*

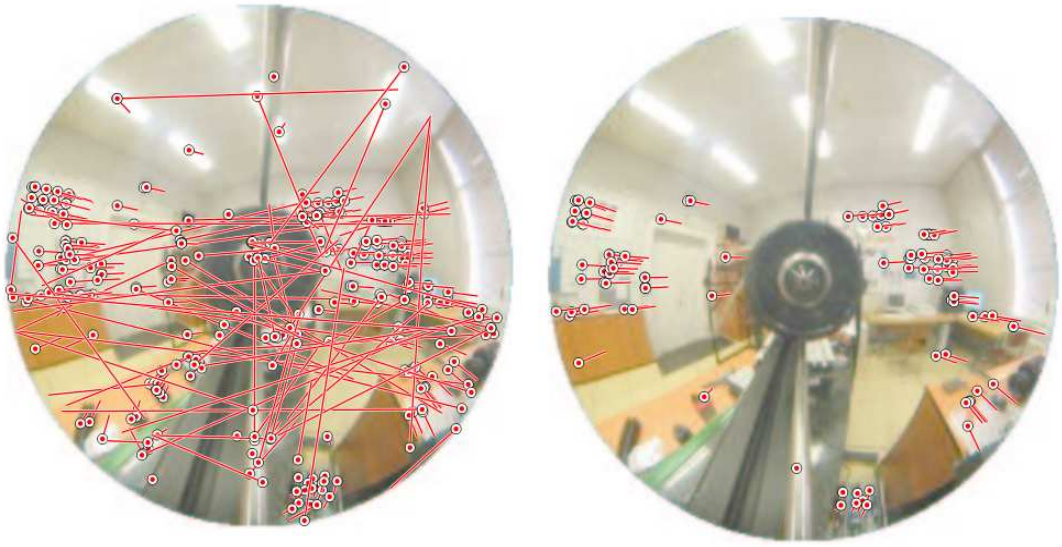


Figure 4.2: *Micusik's method. Example of outliers detection. Left. Tentative correspondences. Right. Validated inliers and their matches in the next image.*

the intersection of the optical axis with the image. They introduce a new representation for images of points and lines in catadioptric images which they call the circle space. This circle space includes imaginary circles, one of which is the image of the absolute conic. They formulate the epipolar constraint in this space and establish a new 4×4 catadioptric fundamental matrix. Then, calibration is performed in two steps : first, they estimate the fundamental matrix $\mathbf{F}$ from manual point correspon-

dences and second, from **F**, they extract the intrinsic parameters via the image of the absolute conic.

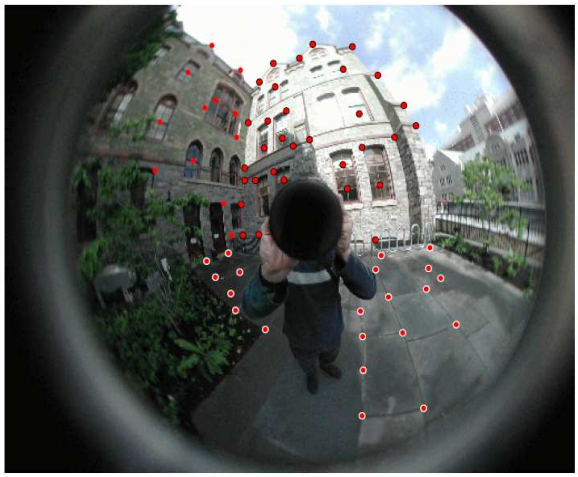


Figure 4.3: *Geyer's method. Points used to calibration purposes.*

Fabrizio *et al.* [FTB02] proposed a calibration method which estimates the intrinsic parameters of the CCD camera and the pose parameters of the CCD camera with respect to the mirror (misalignment). The method is based on the ingenious and simple idea that the mirror's external and internal bounderies can be used as a 3D calibration pattern (See Figure 4.4) . The calibration method is based on the principle used for planar cameras. In this technique, two different 3D planes are required for calibration. Therefore, in this work, the authors use the boundaries of two circles: the circular edge of the upper mirror, and the edge of the intersection of the black needle with the mirror. These two circles lie on two paralel planes which are used by the authors to calibrate the sensor. The proposed algorithm can be applied on most catadioptric sensors if the parameters of the mirror surface are known.

In [BV08], the last reviewed paper, the authors propose a method for automatic calibration of a catadioptric camera. The authors use this kind of sensor in Unmanned Aerial Vehicles (UAV), therefore, the calibration algorithm can be run before mounting the camera on the UAV or even during the flight. The method is against vibrations and shocks. The approach is similar to [YZ05] but it determines, by particle filtering, the most probable calibration parameters with respect to two geometric constraints of lines. First, if the calibration parameters are known, then the conic equation is more constrained and thus the line image can be fitted more accurately. The second constraint is based on the following properties: a line is simply projected as a great circle in the sphere, the great circles associated to parallel lines intersect in two antipodal points in the sphere and the antipodal points (or vanishing points) associated to orthogonal sets of parallel lines are also orthogonal in the sphere. The drawback of this method is the need for estimating the equivalent sphere (i.e. the

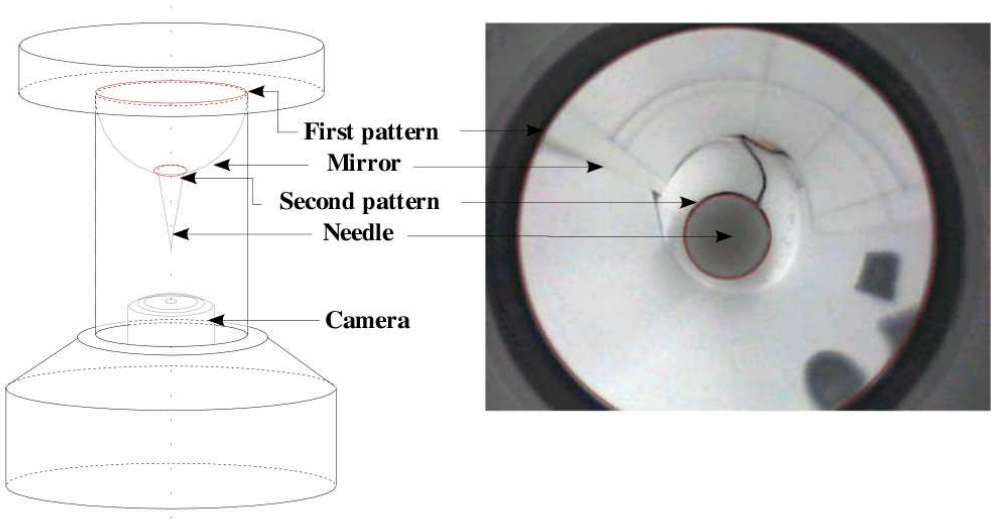


Figure 4.4: *Fabrizio's method. Each of the circles in the image corresponds to a known section of the mirror. They are used as a calibration pattern.*

calibration parameters) that permits one to verify the geometric constraints of the vanishing points associated to parallel and orthogonal lines. Figure 4.5 shows the final result of vanishing point extraction on real data after calibration using the proposed particle filter algorithm.

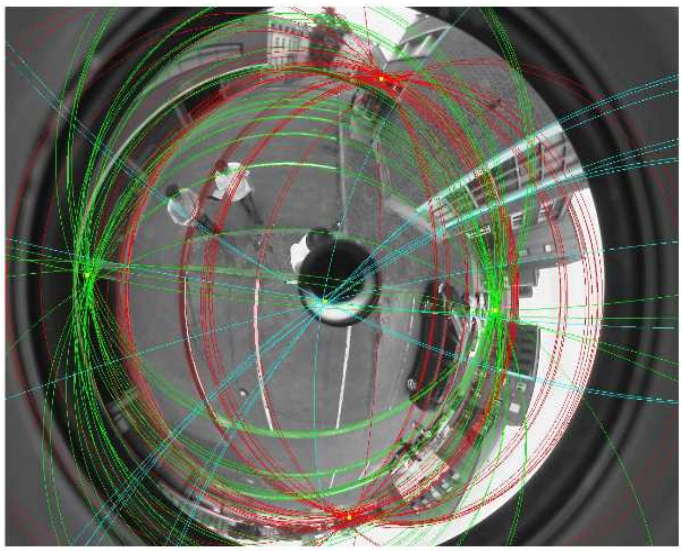


Figure 4.5: *Bazin's method. Lines are used to calibration purposes.*

At this point, we explain the importance and motivation of this thesis. Like traditional calibration techniques, the rewieved self-calibration techniques present some limitations such as dependence on a specific mirror and on the usage of the mirror. Thus, it is not straightforward to implement the same calibration technique for a different mirror. Also, they require feature point correspondences which is not an easy problem because of the high distortions due to the mirror and the repetitive texture in urban enviroments. Therefore, regarding the existing limitations on catadioptric camera self-calibration techniques, we were motivated to propose a new algorithm. The method is based on the tracking of a plane, it stays valid for all central catadioptric systems and needs no prior knowledge about camera, lens or mirror parameters. The only requeriment is to select or detect a plane in the image scene. Then, calibration is done automatically. In the next section we will explain the details of our method.

4.2 Proposed Direct Self-Calibration Method

The contribution of this thesis is aimed at an on-line self-calibration employing several views in a tracked sequence.

From Chapter 3 (Section 3.2.1), we remind: $\boldsymbol{\gamma}$ contains the camera intrinsic parameters $\beta_1, \beta_2, u_0, v_0$, $\boldsymbol{\xi}$ contains the mirror parameter ξ and $\mathbf{z}$ contains the homography parameters $h_1, h_2, h_3, h_4, h_5, h_6, h_7, h_8, h_9$.

Let $\overline{\boldsymbol{\xi}}$, $\overline{\boldsymbol{\gamma}}$ be the true intrinsic parameters of the sensor and $\overline{\mathbf{z}}$ be the homography parameters. The interest is to find the set of parameters $\widehat{\boldsymbol{\xi}}$, $\widehat{\boldsymbol{\gamma}}$ and $\widehat{\mathbf{z}}$, that fits best with the true sensor parameters such that the current image is aligned with the reference template. Therefore, when Equation (4.1) is satisfied, the calibration of the sensor is achieved.

$$I'(\mathbf{w}(\widehat{\mathbf{z}}, \widehat{\boldsymbol{\xi}}, \widehat{\boldsymbol{\gamma}}, \mathbf{p})) = I(\mathbf{p}) \tag{4.1}$$

In order to achieve the goal, our algorithm is based on the work that has been shown in Chapter 3. In Chapter 3, a planar region in the scene was tracked using an uncalibrated central catadioptric camera. The aim was not to recover the true intrinsic parameters but to align the image regions along the sequence with whatever intrinsic and extrinsic parameters that did the job the best. Now, we focus on how to update the parameters. In the experiments of the previous chapter, we have seen that shadows and specularities diminished the performance of the algorithm in terms of quality of the estimates. The reprojection error increases with either shadows or specularities. Therefore, in order to improve the least-square estimates, reprojection errors after alignment are verified. If the reprojection error between two consecutive images is bigger than a certainly threshold, intrinsic parameters are not updated. The set of intrinsic parameters remains unchanged until the moment when a reprojection error smaller than the threshold occurs. Then, the threshold takes the value of the last reprojection error and the intrinsic parameters are updated. This process continues all along the sequence. The final calibration values are obtained

at the end of the sequence.

4.2.1 Calibration parameters

4.2.1.1 Intrinsic parameters

The intrinsic parameters to be recovered are essentially the mirror parameter ξ and these from Equation (3.1) -the generalized focal lengths $\beta_1 = k_u f\eta$ and $\beta_2 = k_v f\eta$, the principal point (u_0, v_0) and the skew s. The skew parameter takes into account the fact that, if the retinal plane of the CCD is not orthogonal to the optical axis of the lens, a CCD element (a pixel) will not be projected as an orthogonal rectangle, but as a parallelogram [Fau93]. Most often, for commercial cameras, we can assume $s = 0$ [Tsa89]. So, it is not necessary to introduce more complications to the model by assuming that the skew is not zero. The aspect ratio ($r = \frac{\beta_1}{\beta_2}$), on the other hand, is not always 1 ($\beta_1 = \beta_2$) as is commonly assumed. Nevertheless, whatever is the true value of the aspect ratio, this is a remarkably stable parameter, and will not change due to zooming or focusing. If the cameras used in the acquisition of the images are known or are available, it is possible to calibrate for this parameter once, and use reliably the value found in any other occasion. For images with unknown origin, this procedure is not possible, and the assumption of the aspect ratio being equal to 1 cannot be blindly used.

4.2.1.2 Extrinsic parameters

We will compute an homography instead of $\mathbf{R}$ and $\mathbf{t}$ directly. Although this approach does not use the minimal amount of parameters and could lead to a less stable estimate, it circumvents the issue of choosing the correct homography decomposition (for each homography, a translation $\mathbf{t}$ up to a scale factor and a rotation $\mathbf{R}$ can be extracted).

4.2.2 Experimental Results

In order to validate the proposed self-calibration method we performed experiments with synthetic data and real data. For the synthetic data we have created an image sequence considering a perspective camera and a parabolic mirror shape. No synthetic experiments with hyperbolic and elliptical mirror shapes have been reported because we obtained results similar to the parabolic mirror shape. Therefore, we considered more important to report experiments with real sensors. For the real data we have tested the algorithm with the two kinds of catadioptric omnidirectional cameras (parabolic and hyperbolic) that we have at INRIA Sophia Antipolis Méditerranée (See Figure 3.6).

4.2.3 Synthetic data

In the following we will show a series of experiments with synthetic data. We have simulated the case of an omndirectional camera using a parabolic mirror.

4.2.3.1 Omnidirectional Parabolic camera

Our experimental setup consists of an image sequence composed of 120 images. To create this sequence we transformed a real parabolic image assuming constant intrinsic parameters such as a catadioptric camera with a parabolic mirror ($\xi = 1$), a generalized focal length $\beta_1 = -250$, $\beta_2 = -250$ and an image center $u_0 = 512$ and $v_0 = 384$. The images were spaced by the homography parameters depicted in Figure 4.6. There is noise due to discretisation. An approximate intensity in the new position is computed by bilinear interpolation.

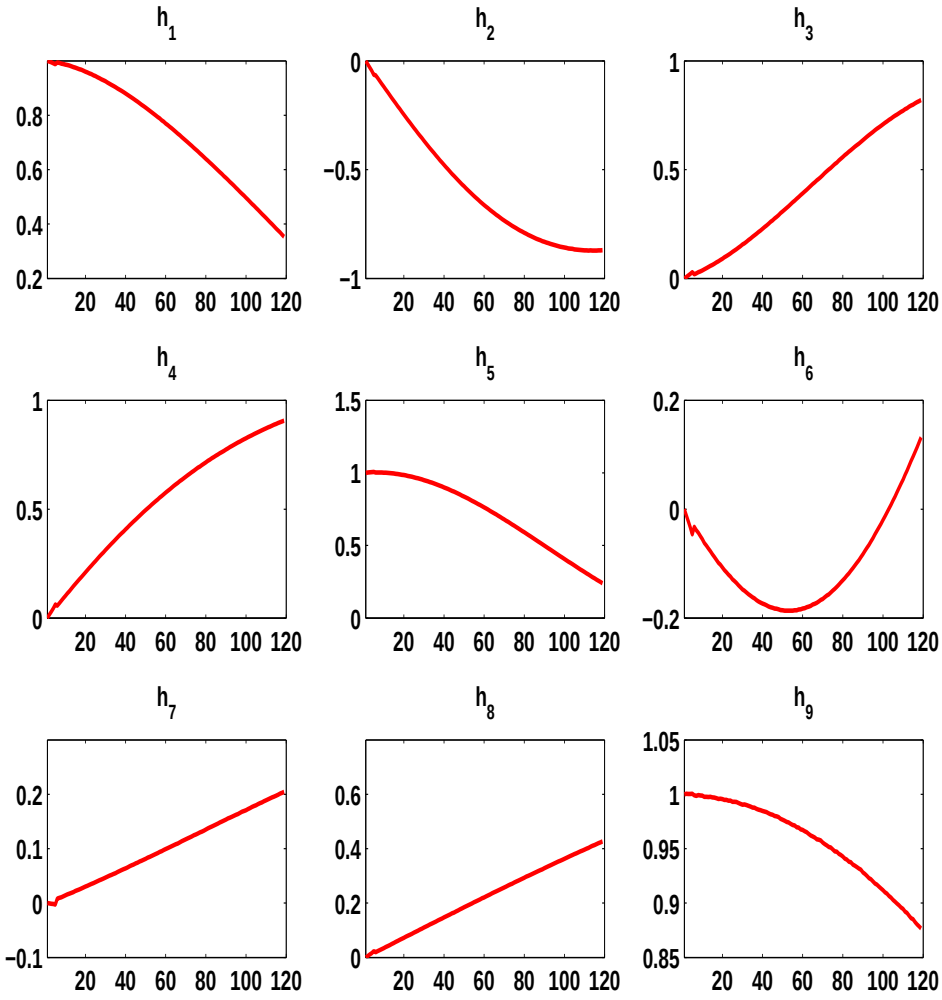


Figure 4.6: *Homography parameters. These parameters were used to create the para-catadioptric image sequence of the self-calibration experiment with synthetic data in Section 4.2.3.1.*

Since the objective is to validate the self-calibration approach, we considered un-

knonwn intrinsic parameters to start the minimization. The initial guess for the intrinsic parameteres was $\xi = 0.8$, $\beta_1 = -125$, $\beta_2 = -125$, $u_0 = 522$ and $v_0 = 394$. The initial guess for the homography parameters was given by the 3×3 identity matrix. Figure 4.7 shows 6 calibration images of the test sequence with the tracked image region marked in red. Figures 4.8 shows the estimated homography parameters during the tracking sequence. Figures 4.9 to 4.13 in this section show the estimated intrinsic parameters. In the beginning of the estimation, the intrinsic parameters are not computed because the homography is the identity. That means that the camera has not moved, therefore, the intrinsic parameters may take any value. To avoid this, we skip a few images to ensure the camera has moved and the intrinsic parameters will take reasonable values that allow a correct convergence. To smooth the noise due to discretisation we applied Kalman filtering to the convergence curve. As we can see the estimation is stabilized after 50 images and subsequent information does not have a strong influence. The final calibration values are obtained by computing the average value of the last 50 estimates. Table 4.1 shows the calibration results for the intrinsic parameters of the sensor.

Table 4.1: *Intrinsic parameters. Calibration results for the experiment with the synthetic para-catadioptric image sequence described in Section 4.2.3.1.*

	ξ	β_1	β_2	u_0	v_0
Ground truth	1.00	-250.00	-250.00	512.00	384.00
Estimated	0.99	-249.99	-249.99	511.99	384.13

4.2.4 Real data

In the following we will show experiments with real data. Two different catadioptric omnidirectional cameras will be considered.

4.2.4.1 Omnidirectional Parabolic camera

The image sequence is composed of 250 images of size 1024×768 combining a camera with a parabolic mirror. To show that the proposed method finds a good aproximation of the true intrinsic parameters, calibration of the sensor was pursued according to [MBMR06]. Determined intrinsic parameters for the real data sequence are $\xi = 1.0$, $\beta_1 = -295$, $\beta_2 = -294$, $u_0 = 519$ and $v_0 = 382$.

Since the objective is to validate the self-calibration approach, we considered the following guess of parameters to star the minimization. The initial guess for the intrinsic parameteres was $\xi = 1.0$, $\beta_1 = -100$, $\beta_2 = -100$, $u_0 = 512$ and $v_0 = 384$. The initial guess for the homography parameters was given by the 3×3 identity matrix. Figure 4.21 shows 6 calibration images of the test sequence with the tracked image region marked in red. Figures 4.22 shows the estimated homography parameters during the tracking sequence. Figures 4.23 to 4.27 in this section show the estimated intrinsic parameters. In the beginning of the estimation, the intrinsic parameters are not computed because the homography is the identity. That means

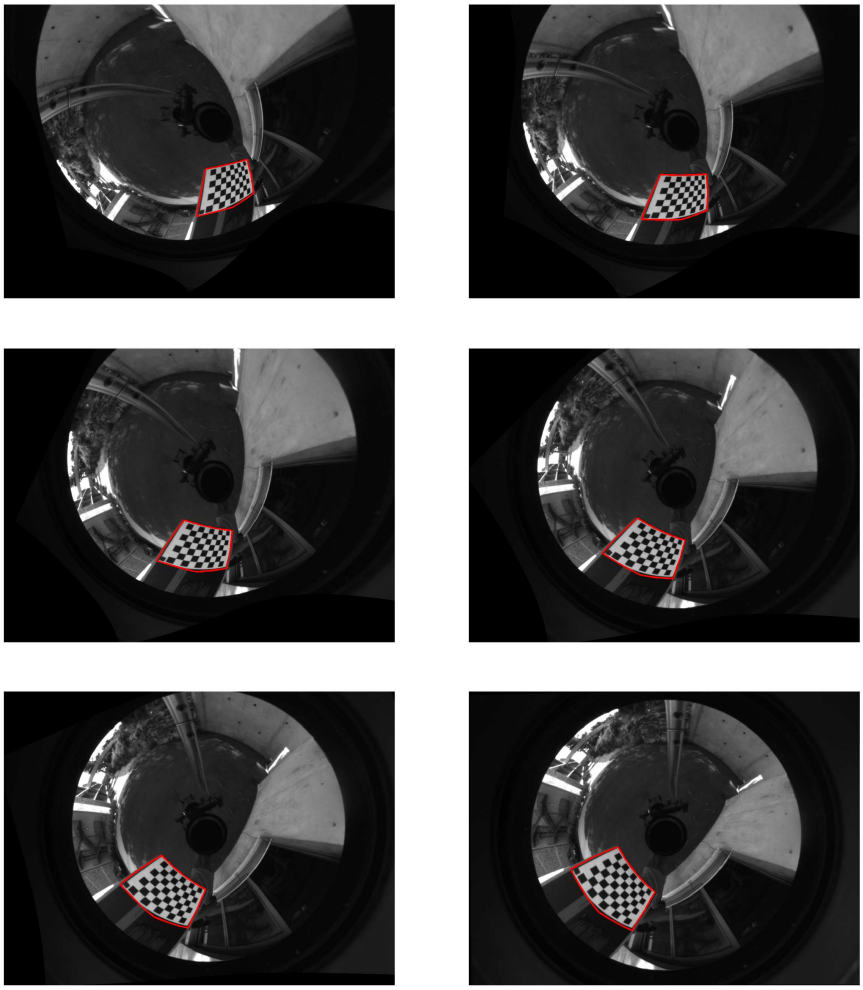


Figure 4.7: *Calibration sequence. Some images of the synthetic calibration sequence described described in Section 4.2.3.1.*

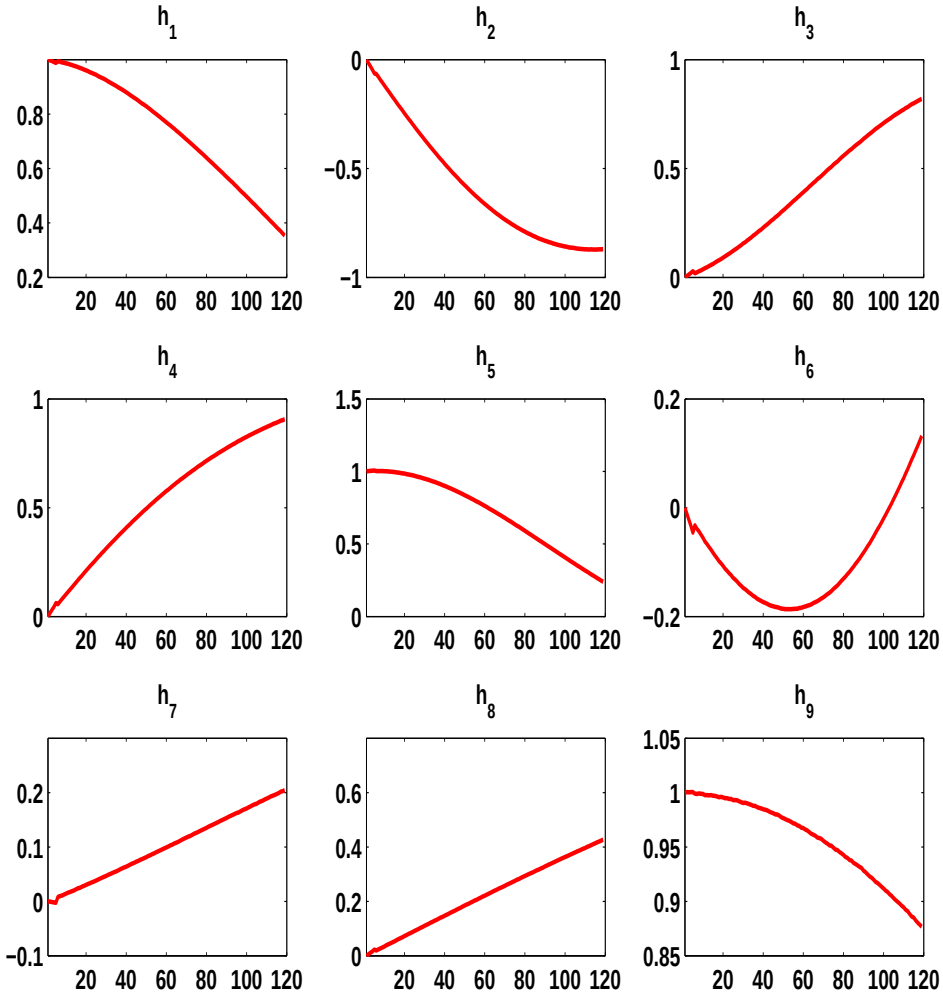


Figure 4.8: *Estimated homography parameters. Parameters obtained after the self-calibration experiment with the synthetic para-catadioptric image sequence described in Section 4.2.3.1.*

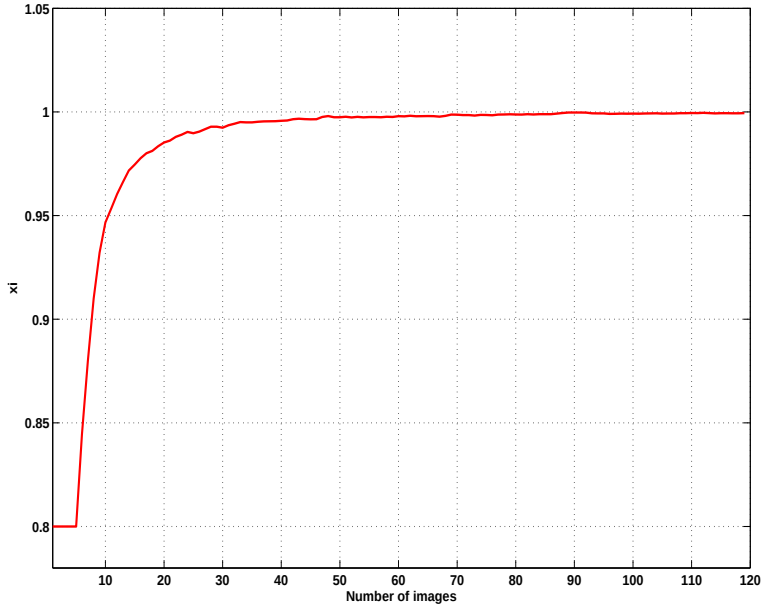


Figure 4.9: *Estimated parameter ξ for the experiment with the synthetic paracatadioptric image sequence described in Section 4.2.3.1. The initial guess has been set to $\xi = 0.8$.*

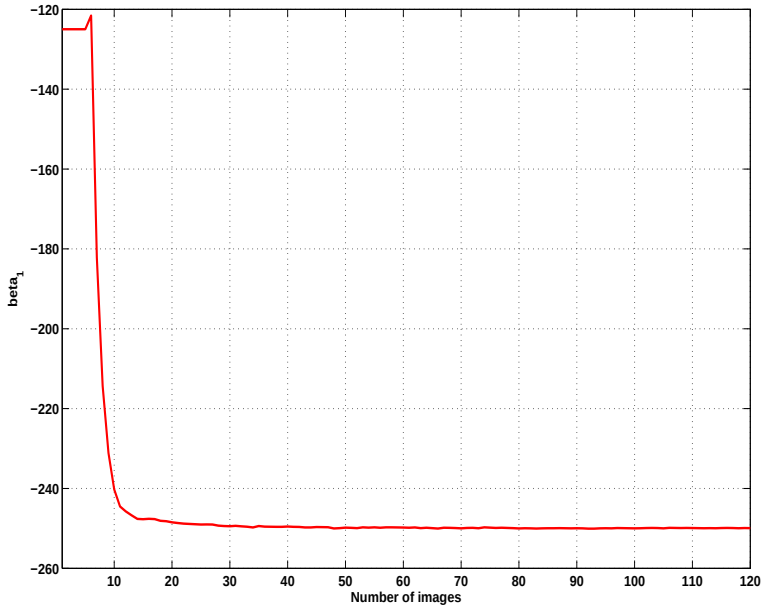


Figure 4.10: *Estimated parameter β_1 for the experiment with the synthetic paracatadioptric image sequence described in Section 4.2.3.1. The initial guess has been set to $\beta_1 = -125$.*

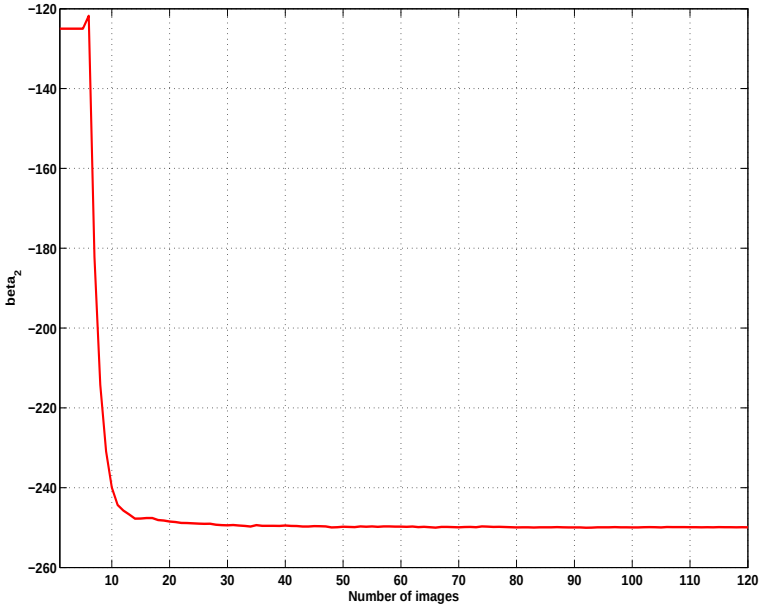


Figure 4.11: *Estimated parameter β_2 for the experiment with the synthetic para-catadioptric image sequence described in Section 4.2.3.1. The initial guess has been set to $\beta_2 = -125$.*

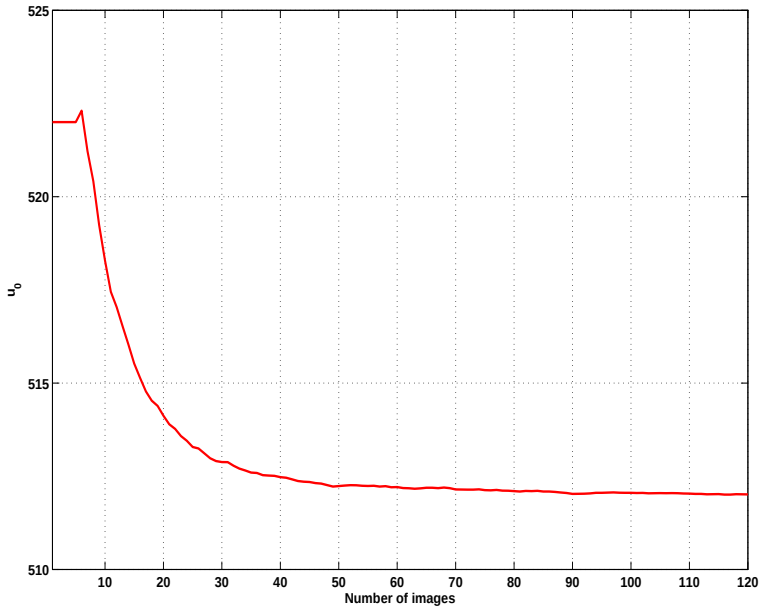


Figure 4.12: *Estimated parameter u_0 for the experiment with the synthetic para-catadioptric image sequence described in Section 4.2.3.1. The initial guess has been set to $u_0 = 522$.*

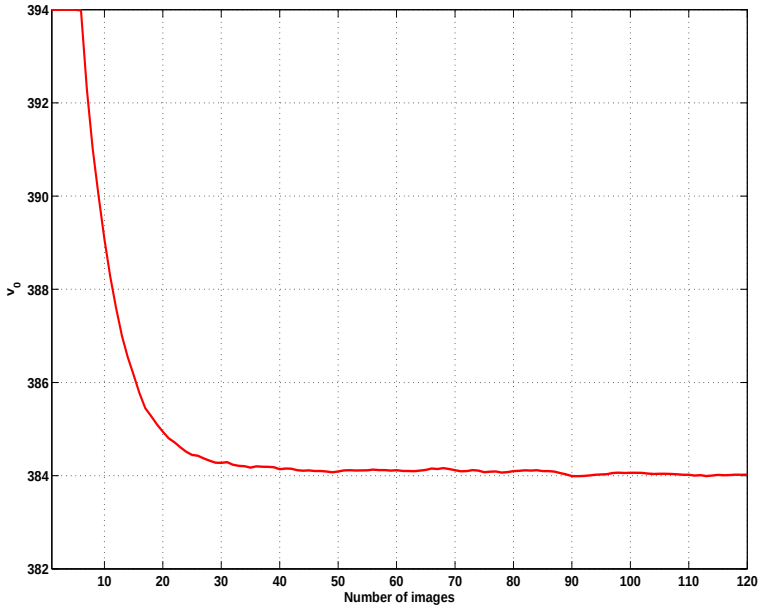


Figure 4.13: *Estimated parameter v_0 for the experiment with the synthetic para-catadioptric image sequence described in Section 4.2.3.1. The initial guess has been set to $v_0 = 394$.*

that the camera has not moved, therefore, the intrinsic parameters may take any value. To avoid this, we skip a few images to ensure the camera has moved and the intrinsic parameters will take reasonable values that allow a correct convergence. To smooth the noise due to discretisation we applied Kalman filtering to the convergence curve. As we can see, the estimation, for some parameteres, stabilizes after 50 images and the subsequent information does not have strong influence. The strong variance of parameters u_0 and v_0 is symptomatic of a sensitive parameter. It may be due to illumination changes and blur effect. From synthetic data we have seen that if noise due to illumination and blur effect are not present the convergence of parameters stabilizes after 50 images, therefore, in order to provide a better estimation of parameters, we considered image sequences with more than 200 images. The final calibration values are obtained by computing the average value of the last 100 estimates. Table 4.2 shows the calibration results for the intrinsic parameters of the sensor.

Table 4.2: *Intrinsic parameters. Calibration results for the experiment with the para-catadioptric image sequence described in Section 4.2.4.1.*

	ξ	β_1	β_2	u_0	v_0
Our method	1.13	-315.32	-313.75	513.88	378.86
[MBMR06]	1.00	-295.00	-294.00	519.00	382.00

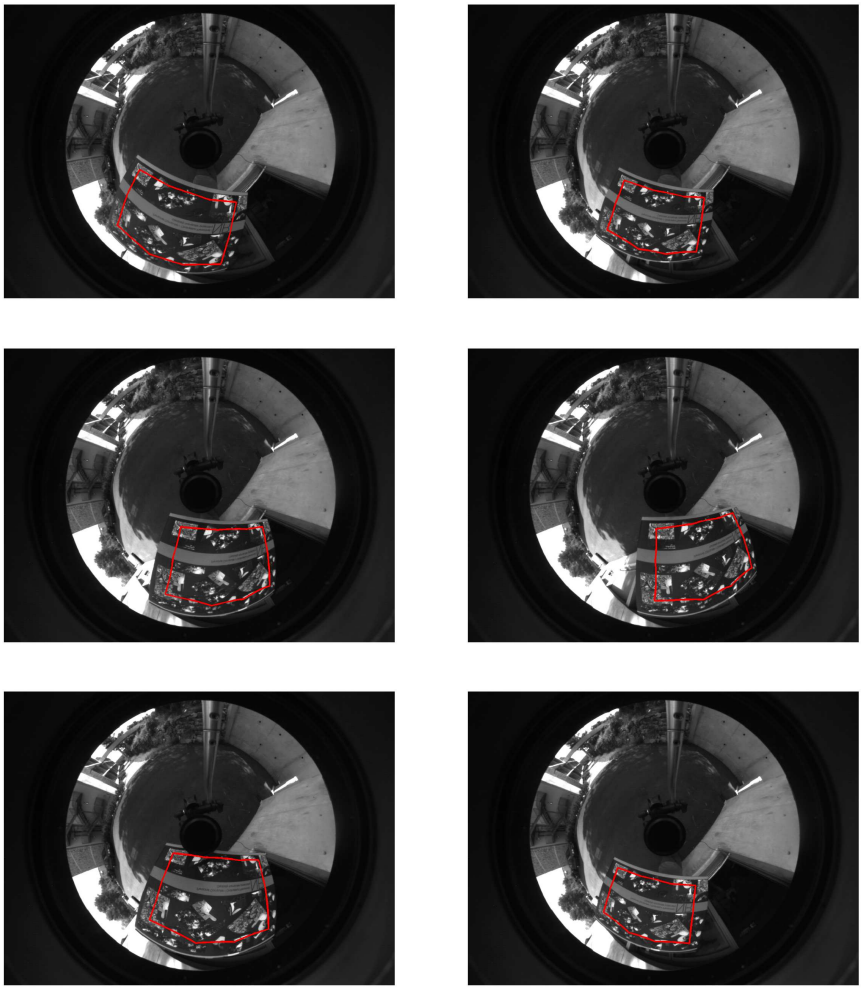


Figure 4.14: *Calibration sequence. Some images of the calibration para-catadioptric image sequence described in Section 4.2.4.1.*

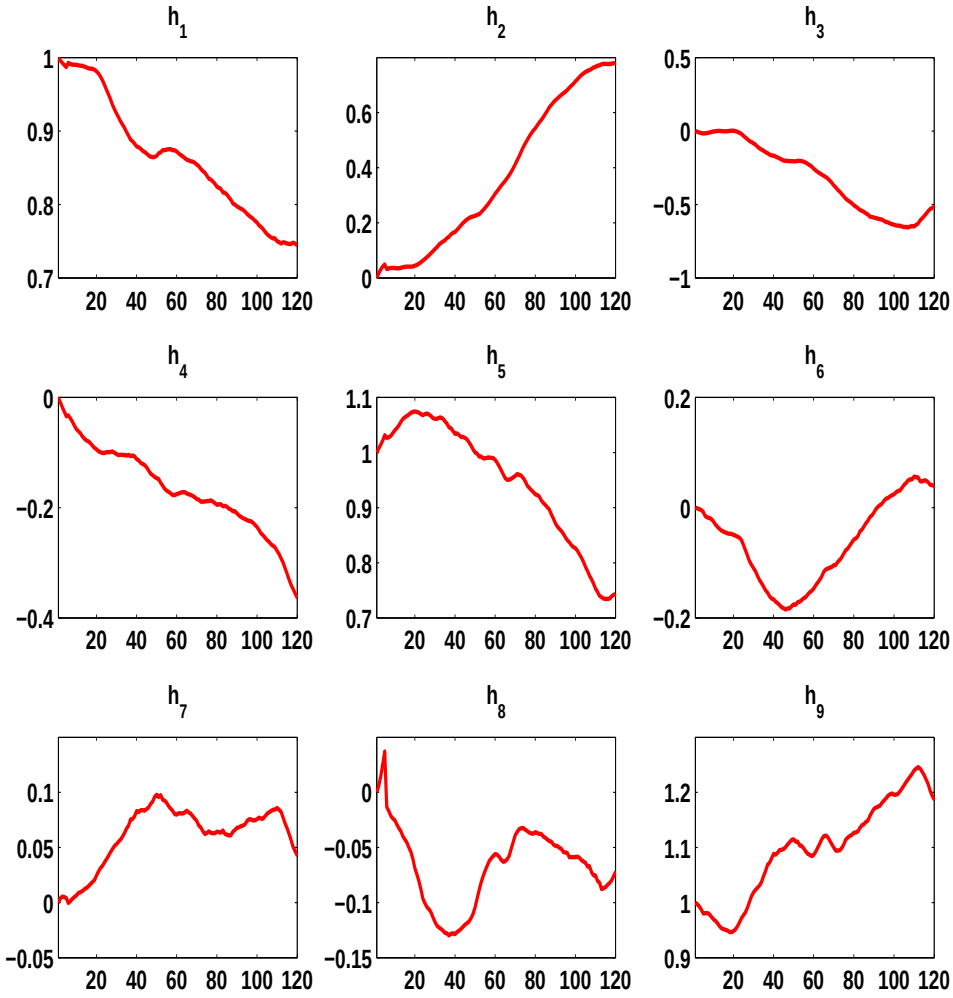


Figure 4.15: *Estimated homography parameters. Parameters obtained after the self-calibration experiment with the para-catadioptric image sequence described in Section 4.2.4.1.*

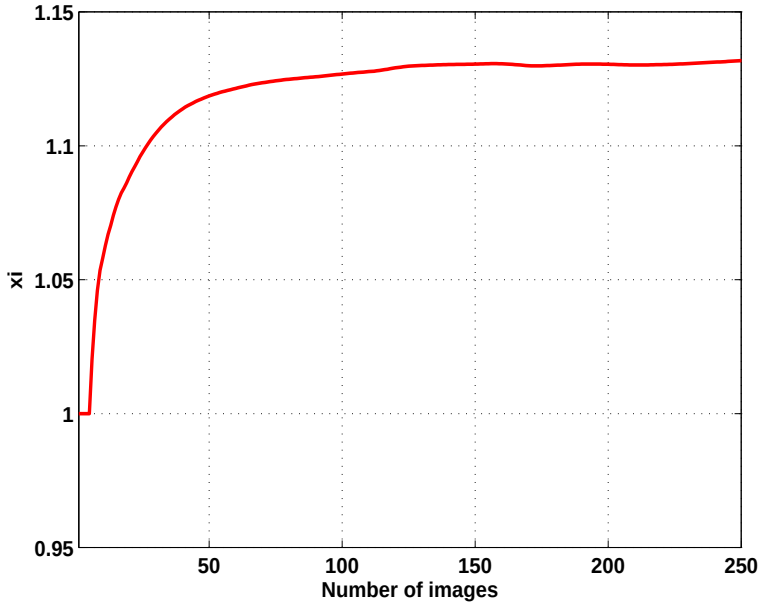


Figure 4.16: *Estimated parameter ξ for the para-catadioptric image sequence described in Section 4.2.4.1. The initial guess has been set to $\xi = 0.8$.*

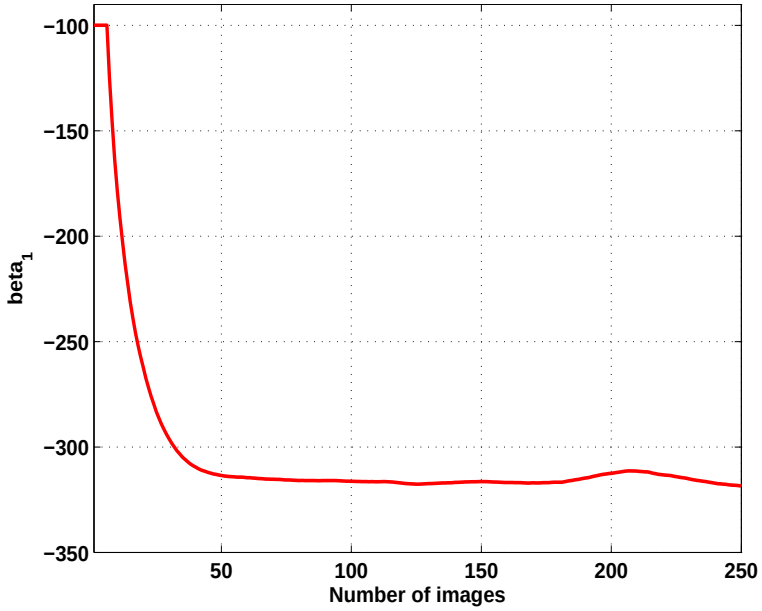


Figure 4.17: *Estimated parameter β_1 for the para-catadioptric image sequence described in Section 4.2.4.1. The initial guess has been set to $\beta_1 = -100$.*

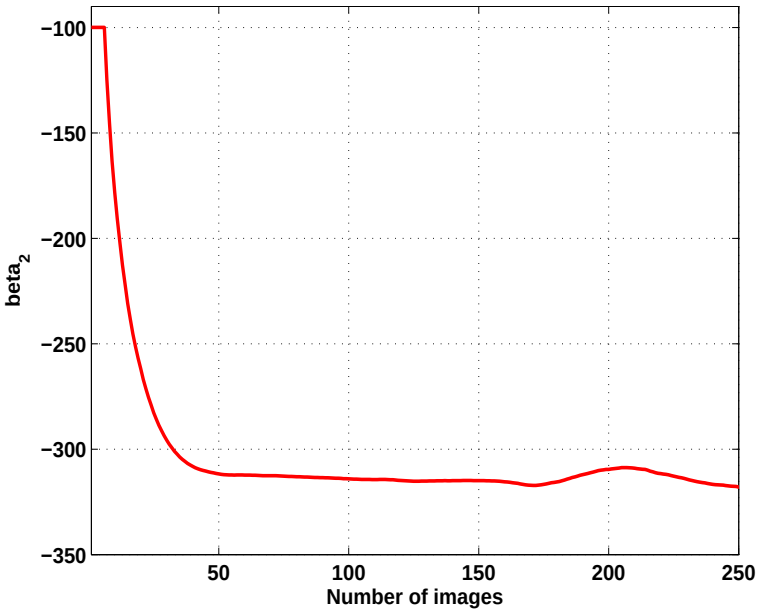


Figure 4.18: *Estimated parameter β_2 for the para-catadioptric image sequence described in Section 4.2.4.1. The initial guess has been set to $\beta_2 = -100$.*

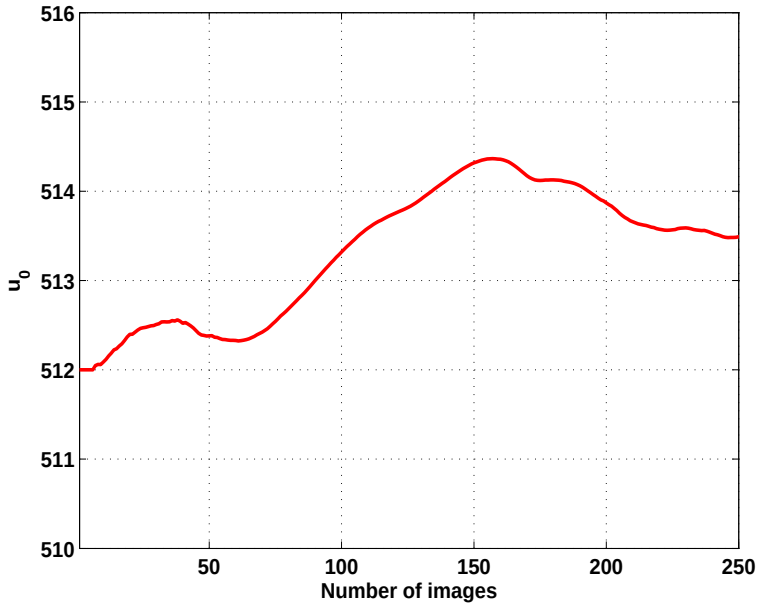


Figure 4.19: *Estimated parameter u_0 for the para-catadioptric image sequence described in Section 4.2.4.1. The initial guess has been set to $u_0 = 512$.*

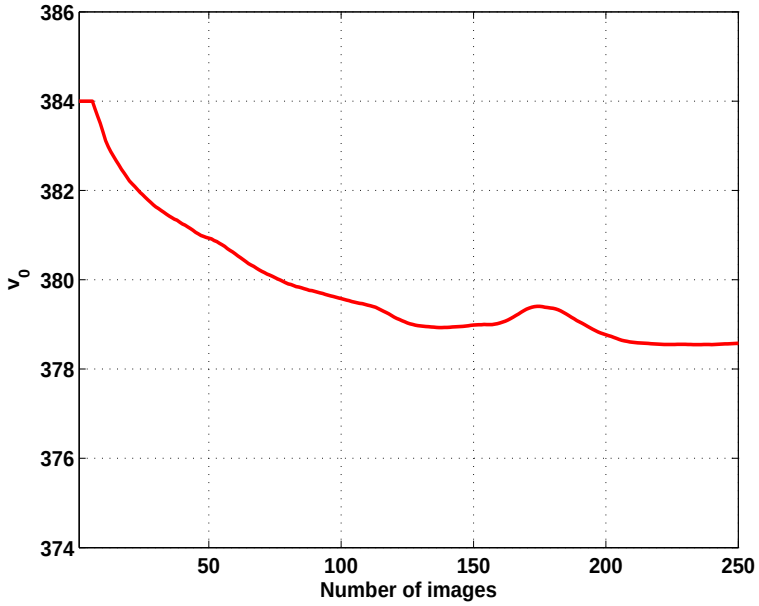


Figure 4.20: *Estimated parameter v_0 for the para-catadioptric image sequence described in Section 4.2.4.1. The initial guess has been set to $v_0 = 384$.*

4.2.4.2 Omnidirectional hyperbolic camera

The image sequence is composed of 250 images of size 640×480 combining a camera with a hyperbolic mirror. To show that the proposed method finds a good aproximation of the true intrinsic parameters, calibration of the sensor was pursued according to [MBMR06]. Determined intrinsic parameters for the real data sequence are $\xi = 0.88$, $\beta_1 = -141.66$, $\beta_2 = -141.82$, $u_0 = 315.77$ and $v_0 = 244.52$.

In order to star the minimization, the initial guess for the intrinsic parameteres was $\xi = 0.9$, $\beta_1 = -100$, $\beta_2 = -100$, $u_0 = 320$ and $v_0 = 220$. The initial guess for the homography parameters was given by the 3×3 identity matrix. Figure 4.21 shows 6 calibration images of the test sequence with the tracked image region marked in red. Figures 4.22 shows the estimated homography parameters during the tracking sequence. Figures 4.23 to 4.27 in this section show the estimated intrinsic parameters. In the beginning of the estimation, the intrinsic parameters are not computed because the homography is the identity. That means that the camera has not moved, therefore, the intrinsic parameters may take any value. To avoid this, we skip a few images to ensure the camera has moved and the intrinsic parameters will take reasonable values that allow a correct convergence. To smooth the noise due to discretisation we applied Kalman filtering to the convergence curve. As we can see the estimate, for some parameteres, stabilize after 150 images and the subsequent information does not have strong influence. The variance of the intrinsic parameters is due to illumination changes (See Figure 4.21). In order to provide a better estimation of parameters, we considered image sequences with more than 200 images. The final calibration values are obtained by computing the average value of the last 100 stimates. Table 4.3 shows calibration results for the intrinsic parameters of the

sensor.

Table 4.3: *Intrinsic parameters. Calibration results for the experiment with the hyper-catadioptric image sequence described in Section 4.2.4.2.*

	ξ	β_1	β_2	u_0	v_0
Our method	0.92	-133.05	-135.10	333.25	225.01
[MBMR06]	0.88	-141.66	-141.82	315.77	244.52

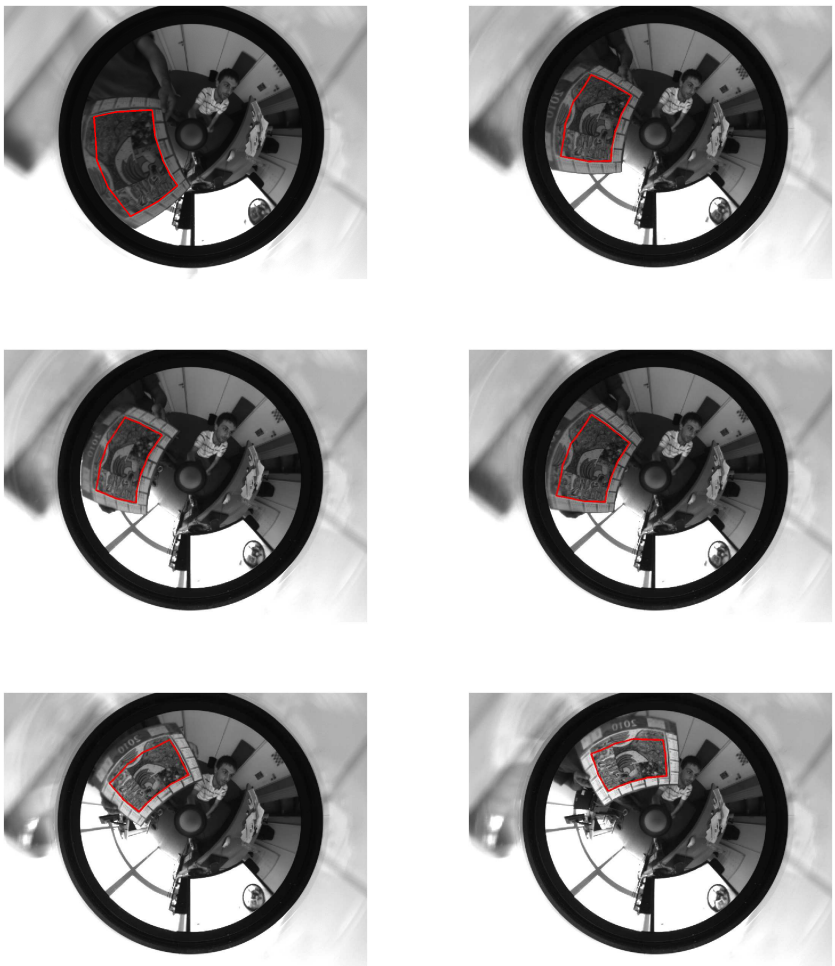


Figure 4.21: *Calibration sequence. Some images of the calibration hyper-catadioptric image sequence described in Section 4.2.4.2.*

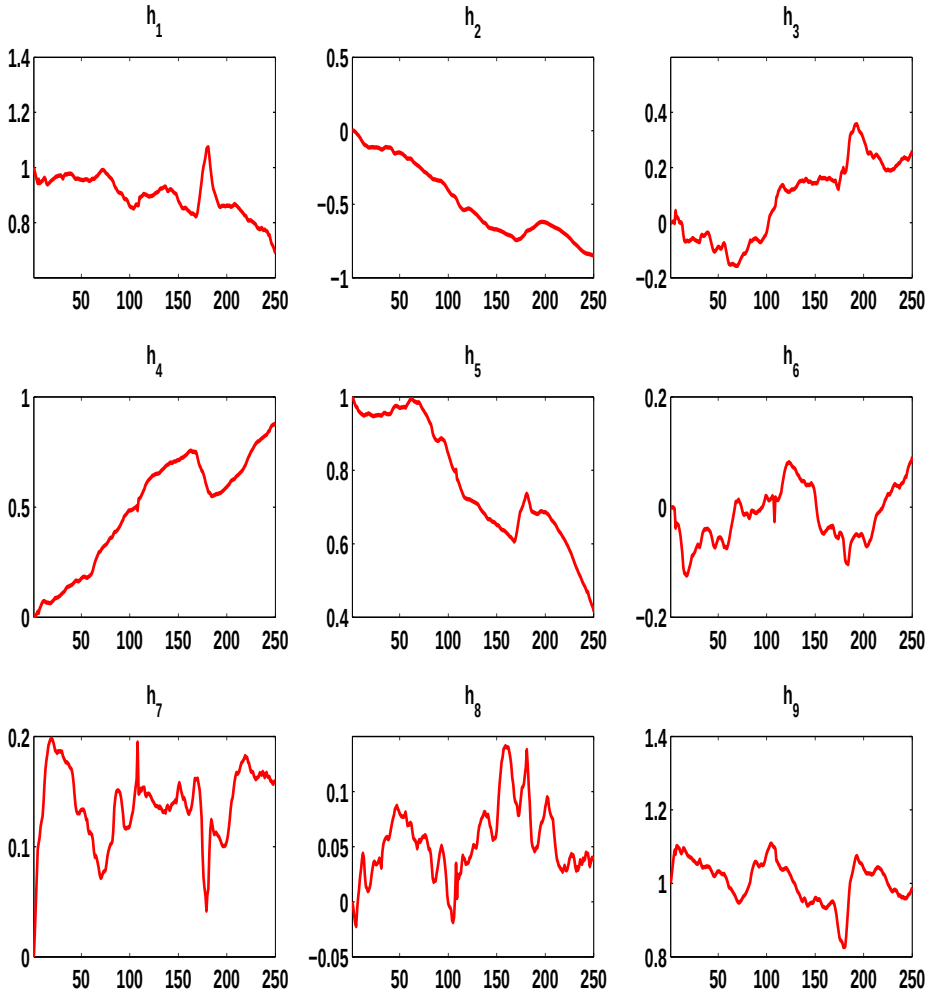


Figure 4.22: *Estimated homography parameters.Parameters obtained after the self-calibration experiment with the para-catadioptric image sequence described in Section 4.2.4.2.*

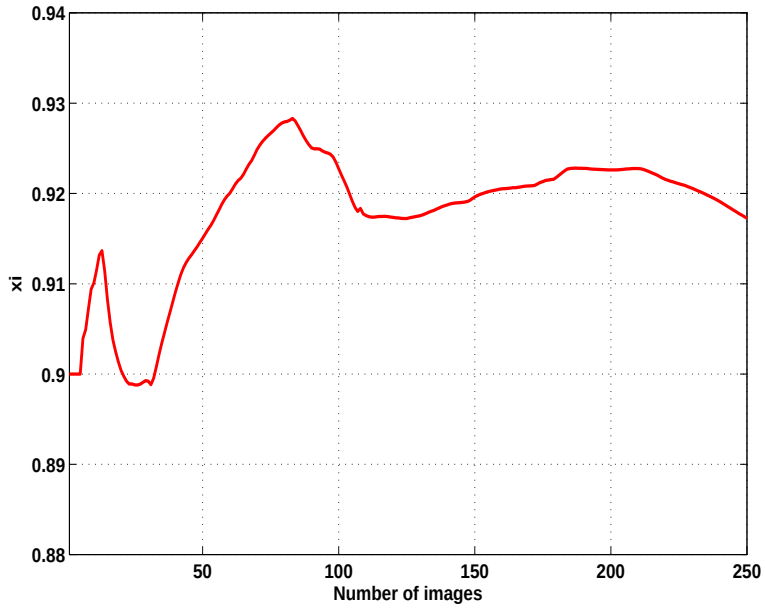


Figure 4.23: *Estimated parameter ξ for the hyper-catadioptric image sequence described in Section 4.2.4.1. The initial guess has been set to $\xi = 0.9$.*

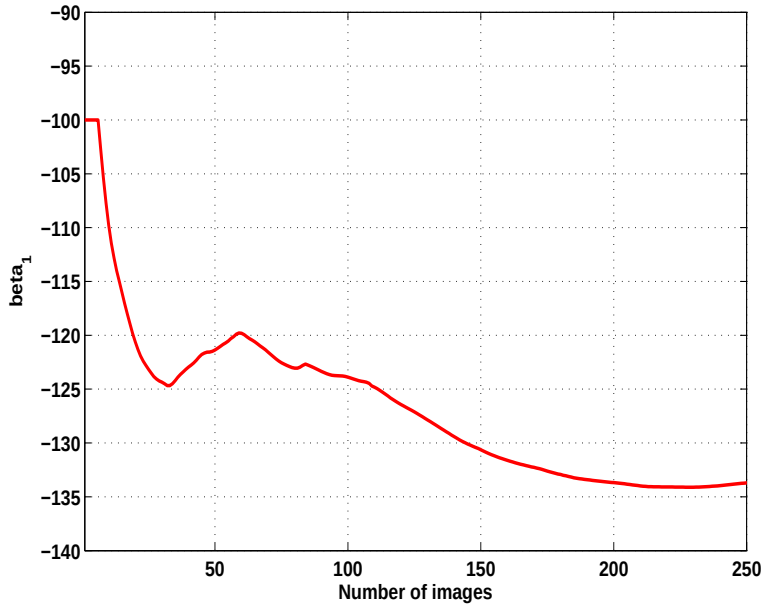


Figure 4.24: *Estimated parameter β_1 for the hyper-catadioptric image sequence described in Section 4.2.4.1. The initial guess has been set to $\beta_1 = -100$.*

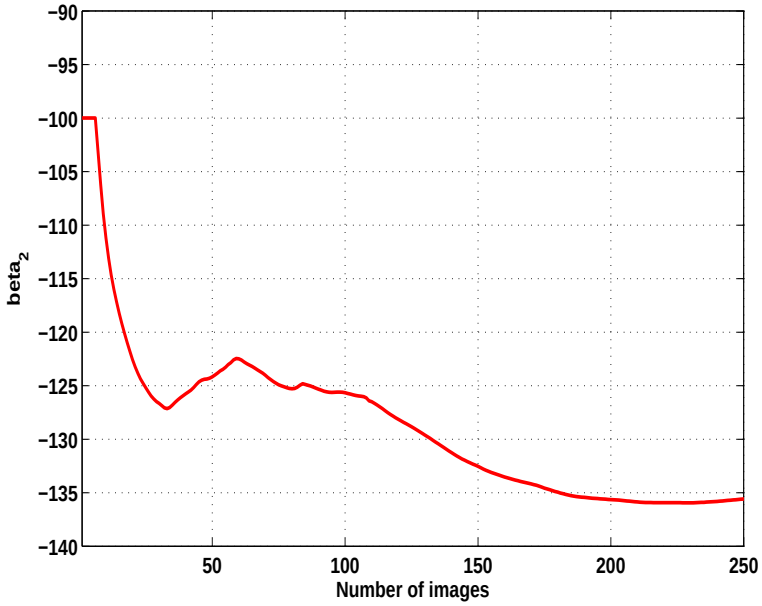


Figure 4.25: *Estimated parameter β_2 for the hyper-catadioptric image sequence described in Section 4.2.4.1. The initial guess has been set to $\beta_2 = -100$.*

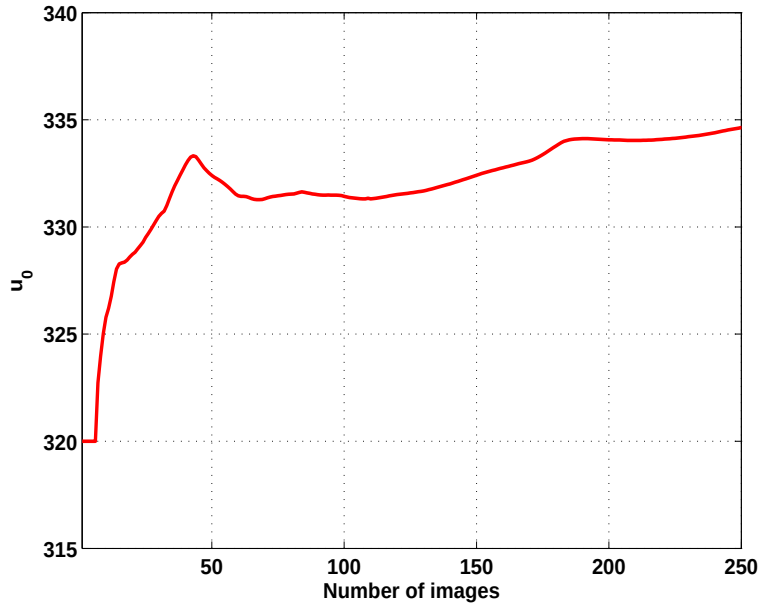


Figure 4.26: *Estimated parameter u_0 for the hyper-catadioptric image sequence described in Section 4.2.4.1. The initial guess has been set to $u_0 = 320$.*

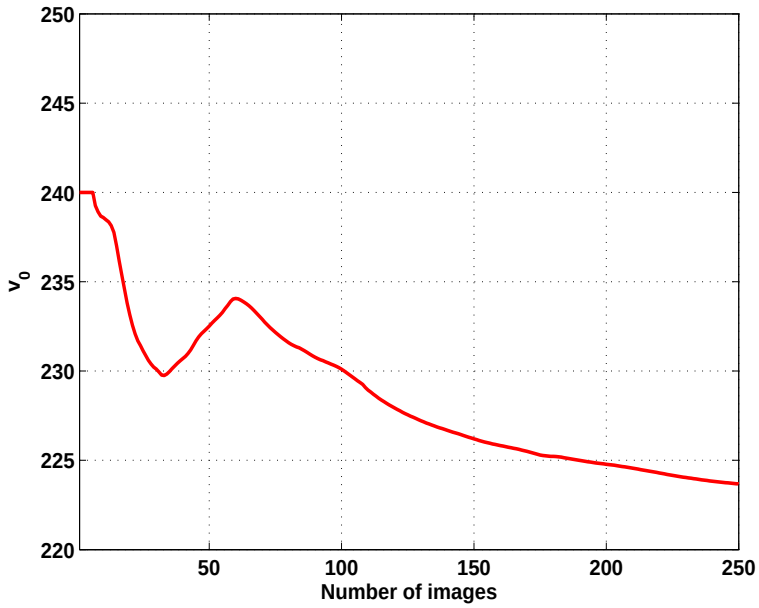


Figure 4.27: *Estimated parameter v_0 for the hyper-catadioptric image sequence described in Section 4.2.4.1. The initial guess has been set to $v_0 = 240$.*

4.3 On the Uniqueness of the Solution for the Calibration of Catadioptric Omndirectional Cameras

Problem statement. In the literature review, different calibration methods have been proposed. Although these approaches have provided appropiate results, none of them has proven the uniqueness of solution for the problem of camera calibration. Therefore, this section aims at providing a proof the uniqueness of solution for the calibration method proposed in this thesis. However, it is worth mentioned that this technique could be extended to different existing methods because the involved parameters are relatively similar.

Let $[\xi, \mathbf{K}, \mathbf{H}]$ and $[\widehat{\xi}, \widehat{\mathbf{K}}, \widehat{\mathbf{H}}]$, with $\xi, \widehat{\xi} \in \mathbb{R}^+$, $\mathbf{H}, \widehat{\mathbf{H}} \in \mathbb{SL}(3)$ and $\mathbf{K}, \widehat{\mathbf{K}}$ upper triangular matrices, be the set of the true parameters and the set of the estimated parameters respectively. Given these two set of parameters, we are interested in verifying that the estimated values converge to the true parameters. The calibration method proposed in this thesis performances calibration as an image registration problem. Thus, the following relation should be verified: $I(\mathbf{w}(\xi, \mathbf{K}, \mathbf{H}, \mathbf{p})) = I(\mathbf{w}(\widehat{\xi}, \widehat{\mathbf{K}}, \widehat{\mathbf{H}}, \mathbf{p})) \ \forall \mathbf{p} \in \mathbb{R}^2$. Considering the assumption made by pixel based calibration algorithms, one has $I^+(I(\mathbf{w}(\mathbf{H}, \xi, \mathbf{K}, \mathbf{p})) = \mathbf{w}(\widehat{\mathbf{H}}, \widehat{\xi}, \widehat{\mathbf{K}}, \mathbf{p})$. Therefore, if this assumption is verified, then Conjecture 1 should be satisfied.

Conjecture 1. Let $[\xi, \mathbf{K}, \mathbf{H}]$ be the set of true parameters and $[\widehat{\xi}, \widehat{\mathbf{K}}, \widehat{\mathbf{H}}]$ be the set of estimated parameters generated by our calibration algorithm; where $0 < \xi, \widehat{\xi} \leq 1$, $\mathbf{K}, \widehat{\mathbf{K}}$ are invertible upper triangular matrices and $\mathbf{H}, \widehat{\mathbf{H}} \in \mathbb{SL}(3) \backslash \mathbb{SO}(3)$. Then, the following condition

$$\mathbf{w}(\mathbf{H}, \xi, \mathbf{K}, \mathbf{p}) = \mathbf{w}(\widehat{\mathbf{H}}, \widehat{\xi}, \widehat{\mathbf{K}}, \mathbf{p}) \, , \forall \mathbf{p} \in \mathbb{R}^2 \tag{4.2}$$

is satisfied if and only if $\widehat{\xi} = \xi$, $\widehat{\mathbf{K}} = \mathbf{K}$ and $\widehat{\mathbf{H}} = \mathbf{H}$.

Remark. Our calibration method uses a unified model that is valid for all central catadioptric omnidirectional cameras. It can be aproximated to traditional cameras when a planar mirror is considered, $\xi = 0$. However, Conjecture 1 does not include $\xi = 0$ and consequently $\widehat{\xi} = 0$ because when $\xi = \widehat{\xi} = 0$ one has exactly the case of a perspective camera with an extra normalization of the point on the normalised plane. As it is well known, there are different solution for the homography matrix. It can be written as $\widehat{\mathbf{H}} = \mathbf{M}\mathbf{H}\mathbf{M}^{-1}$. Therefore the case $\xi = \widehat{\xi} = 0$ does not have an unique solution.

In order to verify Condition (4.23), in what follows we proceed the warping of a point $\mathbf{p}$ using the two sets of parameters (true parameteres of the sensor and the estimated parameters after calibration process) detailed hereafter. Thus, let us introduce some definitions and properties:

- $\widehat{\mathbf{q}} \triangleq (\hat{x}, \hat{y}, 1)^\top$ where $\widehat{\mathbf{q}} = \widehat{\mathbf{K}}^{-1}\mathbf{p} = \widehat{\mathbf{K}}^{-1}\mathbf{I}_3\mathbf{p} = \widehat{\mathbf{K}}^{-1}\mathbf{K}\mathbf{K}^{-1}\mathbf{p} = \mathbf{M}\mathbf{q}$ with $\mathbf{M} \triangleq \widehat{\mathbf{K}}^{-1}\mathbf{K}$. One easily verify that $\mathbf{M}$ is also an upper triangular matrix, and let us denote

$$\mathbf{M} \triangleq \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ 0 & m_{22} & m_{23} \\ 0 & 0 & 1 \end{pmatrix}. \tag{4.3}$$

  where $m_{11} > 0$ and $m_{22} > 0$ because $m_{11} = \frac{k_u}{\widehat{k}_u}$ and $m_{22} = \frac{k_v}{\widehat{k}_v}$. The calibration method takes into account the sign of this parameters, that means $sign(k_u) = sign(\widehat{k}_u)$ and $sign(k_v) = sign(\widehat{k}_v)$. Therefore, m_{11} and m_{22} are positive.

- $\mathbf{H}$ is a true homography matrix represented by

$$\mathbf{H} = \begin{pmatrix} \mathbf{h}_1^\top \\ \mathbf{h}_2^\top \\ \mathbf{h}_3^\top \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix}. \tag{4.4}$$

- $\widehat{\mathbf{H}}$ is an estimated homography matrix represented by

$$\widehat{\mathbf{H}} = \begin{pmatrix} \widehat{\mathbf{h}}_1^\top \\ \widehat{\mathbf{h}}_2^\top \\ \widehat{\mathbf{h}}_3^\top \end{pmatrix} = \begin{pmatrix} \widehat{h}_{11} & \widehat{h}_{12} & \widehat{h}_{13} \\ \widehat{h}_{21} & \widehat{h}_{22} & \widehat{h}_{23} \\ \widehat{h}_{31} & \widehat{h}_{32} & \widehat{h}_{33} \end{pmatrix}. \tag{4.5}$$

In this thesis we set $\det(\mathbf{H}) = 1$ in order to fix the scale factor of the homography. Therefore, throughout this proof we respect the constraint that $\det(\mathbf{H}) = \det(\widehat{\mathbf{H}}) = 1$. This constraint also indicates that $\mathbf{h}_i \neq 0$ and $\widehat{\mathbf{h}}_i \neq 0$ for all $i \in \{1, 2, 3\}$.

4.3.1 Warping using the true parameters.

The warping function relates every point $\mathbf{p}$ of an image frame ($\mathbf{p}$ belongs to a planar region on the scene) with a point $\mathbf{p}'$ in a consecutive image frame as follows

$$\mathbf{p}' = \mathbf{w}(\mathbf{H}, \xi, \mathbf{K}, \mathbf{p}) = \mathbf{K} \begin{pmatrix} \frac{\mathbf{h}_1^\top \mathbf{X}_s}{\mathbf{h}_3^\top \mathbf{X}_s - \xi \|\mathbf{H}\mathbf{X}_s\|} \\ \frac{\mathbf{h}_2^\top \mathbf{X}_s}{\mathbf{h}_3^\top \mathbf{X}_s - \xi \|\mathbf{H}\mathbf{X}_s\|} \\ 1 \end{pmatrix} \tag{4.6}$$

where

$$\mathbf{X}_s = \begin{pmatrix} \alpha\, x \\ \alpha\, y \\ \alpha - \xi \end{pmatrix} \tag{4.7}$$

and

$$\alpha = \frac{\xi + \sqrt{1 + (1 - \xi^2)(x^2 + y^2)}}{x^2 + y^2 + 1} \tag{4.8}$$

From (4.7) and using the true parameters of the sensor one obtains

$$\mathbf{X}_s = \frac{\alpha}{\xi + \sqrt{A_2}}\, \overline{\mathbf{X}}_s \tag{4.9}$$

with

$$\overline{\mathbf{X}}_s = \xi \begin{pmatrix} x \\ y \\ -x^2 - y^2 \end{pmatrix} + \sqrt{A_2} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{4.10}$$

and

$$A_2 = 1 + (1 - \xi^2)(x^2 + y^2) \tag{4.11}$$

Replacing (4.9) into (4.6) one obtains the coordinates of the warped point $\mathbf{p}'$ as follows

$$\mathbf{p}' = \mathbf{K} \begin{pmatrix} \frac{\mathbf{h}_1^\top \overline{\mathbf{X}}_s}{\mathbf{h}_3^\top \overline{\mathbf{X}}_s - \xi \|\mathbf{H}\overline{\mathbf{X}}_s\|} \\ \frac{\mathbf{h}_2^\top \overline{\mathbf{X}}_s}{\mathbf{h}_3^\top \overline{\mathbf{X}}_s - \xi \|\mathbf{H}\overline{\mathbf{X}}_s\|} \\ 1 \end{pmatrix} \tag{4.12}$$

4.3.2 Warping using the estimated parameters.

Now, we proceed the warping of the point $\mathbf{p}$ using the estimated parameters. From (4.7) one obtains

$$\widehat{\mathbf{X}}_s = \begin{pmatrix} \widehat{\alpha}\,(m_{11}x + m_{12}y + m_{13}) \\ \widehat{\alpha}\,(m_{22}y + m_{23}) \\ \widehat{\alpha} - \widehat{\xi} \end{pmatrix} \tag{4.13}$$

$$\widehat{\mathbf{X}}_s = \frac{\widehat{\alpha}}{\widehat{\xi} + \sqrt{\widehat{A}_2}}\widehat{\overline{\mathbf{X}}}_s \tag{4.14}$$

with

$$\widehat{\alpha} = \frac{\widehat{\xi} + \sqrt{1 + (1 - \widehat{\xi}^2)((m_{11}x + m_{12}y + m_{13})^2 + (m_{22}y + m_{23})^2)}}{(m_{11}x + m_{12}y + m_{13})^2 + (m_{22}y + m_{23})^2 + 1}$$

$$\widehat{\overline{\mathbf{X}}}_s = \widehat{\xi}\begin{pmatrix} m_{11}x + m_{12}y + m_{13} \\ m_{22}y + m_{23} \\ -(m_{11}x + m_{12}y + m_{13})^2 - (m_{22}y + m_{23})^2 \end{pmatrix} + \sqrt{\widehat{A}_2}\begin{pmatrix} m_{11}x + m_{12}y + m_{13} \\ m_{22}y + m_{23} \\ 1 \end{pmatrix} \tag{4.15}$$

and

$$\widehat{A}_2 = 1 + (1 - \widehat{\xi}^2)((m_{11}x + m_{12}y + m_{13})^2 + (m_{22}y + m_{23})^2) \tag{4.16}$$

Again, replacing (4.14) into (4.6) one obtains the warping of the point $\mathbf{p}$ as follows

$$\widehat{\mathbf{p}}' = \widehat{\mathbf{K}}\begin{pmatrix} \frac{\widehat{\mathbf{h}}_1^\top \widehat{\overline{\mathbf{X}}}_s}{\widehat{\mathbf{h}}_3^\top \widehat{\overline{\mathbf{X}}}_s - \widehat{\xi}\|\widehat{\mathbf{H}}\widehat{\overline{\mathbf{X}}}_s\|} \\ \frac{\widehat{\mathbf{h}}_2^\top \widehat{\overline{\mathbf{X}}}_s}{\widehat{\mathbf{h}}_3^\top \widehat{\overline{\mathbf{X}}}_s - \widehat{\xi}\|\widehat{\mathbf{H}}\widehat{\overline{\mathbf{X}}}_s\|} \\ 1 \end{pmatrix} \tag{4.17}$$

Now, let us look for conditions on $\widehat{\xi}$, $\widehat{\mathbf{K}}$ and $\widehat{\mathbf{H}}$ so that (4.23) is satisfied. From (4.23), (4.12), (4.17) and the definition of $\mathbf{w}(\bullet)$ one deduces

$$\widehat{\mathbf{K}}\begin{pmatrix} \frac{\widehat{\mathbf{h}}_1^\top \widehat{\overline{\mathbf{X}}}_s}{\widehat{\mathbf{h}}_3^\top \widehat{\overline{\mathbf{X}}}_s - \widehat{\xi}\|\widehat{\mathbf{H}}\widehat{\overline{\mathbf{X}}}_s\|} \\ \frac{\widehat{\mathbf{h}}_2^\top \widehat{\overline{\mathbf{X}}}_s}{\widehat{\mathbf{h}}_3^\top \widehat{\overline{\mathbf{X}}}_s - \widehat{\xi}\|\widehat{\mathbf{H}}\widehat{\overline{\mathbf{X}}}_s\|} \\ 1 \end{pmatrix} = \mathbf{K}\begin{pmatrix} \frac{\mathbf{h}_1^\top \overline{\mathbf{X}}_s}{\mathbf{h}_3^\top \overline{\mathbf{X}}_s - \xi\|\mathbf{H}\overline{\mathbf{X}}_s\|} \\ \frac{\mathbf{h}_2^\top \overline{\mathbf{X}}_s}{\mathbf{h}_3^\top \overline{\mathbf{X}}_s - \xi\|\mathbf{H}\overline{\mathbf{X}}_s\|} \\ 1 \end{pmatrix} \tag{4.18}$$

Since $\widehat{\mathbf{K}}$ is invertible, multiplying both sides of (4.18) by $\widehat{\mathbf{K}}^{-1}$ one obtains

$$\begin{pmatrix} \frac{\widehat{\mathbf{h}}_1^\top \widehat{\overline{\mathbf{X}}}_s}{\widehat{\mathbf{h}}_3^\top \widehat{\overline{\mathbf{X}}}_s - \widehat{\xi}\|\widehat{\mathbf{H}}\widehat{\overline{\mathbf{X}}}_s\|} \\ \\ \frac{\widehat{\mathbf{h}}_2^\top \widehat{\overline{\mathbf{X}}}_s}{\widehat{\mathbf{h}}_3^\top \widehat{\overline{\mathbf{X}}}_s - \widehat{\xi}\|\widehat{\mathbf{H}}\widehat{\overline{\mathbf{X}}}_s\|} \\ \\ 1 \end{pmatrix} = \widehat{\mathbf{K}}^{-1}\mathbf{K} \begin{pmatrix} \frac{\mathbf{h}_1^\top \overline{\mathbf{X}}_s}{\mathbf{h}_3^\top \overline{\mathbf{X}}_s - \xi\|\mathbf{H}\overline{\mathbf{X}}_s\|} \\ \\ \frac{\mathbf{h}_2^\top \overline{\mathbf{X}}_s}{\mathbf{h}_3^\top \overline{\mathbf{X}}_s - \xi\|\mathbf{H}\overline{\mathbf{X}}_s\|} \\ \\ 1 \end{pmatrix} \tag{4.19}$$

Replacing (4.3) into (4.19), one obtains

$$\begin{pmatrix} \frac{\widehat{\mathbf{h}}_1^\top \widehat{\overline{\mathbf{X}}}_s}{\widehat{\mathbf{h}}_3^\top \widehat{\overline{\mathbf{X}}}_s - \widehat{\xi}\|\widehat{\mathbf{H}}\widehat{\overline{\mathbf{X}}}_s\|} \\ \\ \frac{\widehat{\mathbf{h}}_2^\top \widehat{\overline{\mathbf{X}}}_s}{\widehat{\mathbf{h}}_3^\top \widehat{\overline{\mathbf{X}}}_s - \widehat{\xi}\|\widehat{\mathbf{H}}\widehat{\overline{\mathbf{X}}}_s\|} \\ \\ 1 \end{pmatrix} = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ 0 & m_{22} & m_{23} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{\mathbf{h}_1^\top \overline{\mathbf{X}}_s}{\mathbf{h}_3^\top \overline{\mathbf{X}}_s - \xi\|\mathbf{H}\overline{\mathbf{X}}_s\|} \\ \\ \frac{\mathbf{h}_2^\top \overline{\mathbf{X}}_s}{\mathbf{h}_3^\top \overline{\mathbf{X}}_s - \xi\|\mathbf{H}\overline{\mathbf{X}}_s\|} \\ \\ 1 \end{pmatrix} \tag{4.20}$$

which implies that

$$\frac{\widehat{\mathbf{h}}_1^\top \widehat{\overline{\mathbf{X}}}_s}{\widehat{\mathbf{h}}_3^\top \widehat{\overline{\mathbf{X}}}_s - \widehat{\xi}\|\widehat{\mathbf{H}}\widehat{\overline{\mathbf{X}}}_s\|} = m_{11}\frac{\mathbf{h}_1^\top \overline{\mathbf{X}}_s}{\mathbf{h}_3^\top \overline{\mathbf{X}}_s - \xi\|\mathbf{H}\overline{\mathbf{X}}_s\|} + m_{12}\frac{\mathbf{h}_2^\top \overline{\mathbf{X}}_s}{\mathbf{h}_3^\top \overline{\mathbf{X}}_s - \xi\|\mathbf{H}\overline{\mathbf{X}}_s\|} + m_{13} \tag{4.21}$$

$$\frac{\widehat{\mathbf{h}}_2^\top \widehat{\overline{\mathbf{X}}}_s}{\widehat{\mathbf{h}}_3^\top \widehat{\overline{\mathbf{X}}}_s - \widehat{\xi}\|\widehat{\mathbf{H}}\widehat{\overline{\mathbf{X}}}_s\|} = m_{22}\frac{\mathbf{h}_2^\top \overline{\mathbf{X}}_s}{\mathbf{h}_3^\top \overline{\mathbf{X}}_s - \xi\|\mathbf{H}\overline{\mathbf{X}}_s\|} + m_{23} \tag{4.22}$$

In order to prove Conjecture 1, we have to prove that (4.21) and (4.22) have a unique solution for $\widehat{\xi}$, $\widehat{\mathbf{K}}$ and $\widehat{\mathbf{H}}$. Due to the complexity of the problem, let us for instance consider a simplified version for which ξ and $\widehat{\xi}$ are known in advance and is equal to 1 (parabolic mirror). More precisely, supposed that the estimated mirror parameter $\widehat{\xi}$ is identically to the true mirror parameter ξ. Thus, we only consider $\widehat{\mathbf{K}}$ and $\widehat{\mathbf{H}}$ as variables for equations (4.21) and (4.22).

Let us analys for $\widehat{\xi} = \widehat{\xi} = 1$ (parabolic mirror).

4.3.3 Parabolic mirror

In this section we will study the uniqueness of the solution for the calibration problem of central catadioptric cameras when a parabolic mirror is used.

Theorem 1 *Let* $[\mathbf{K}, \mathbf{H}]$ *be the set of true parameters and* $[\widehat{\mathbf{K}}, \widehat{\mathbf{H}}]$ *be the set of estimated parameters generated by the calibration algorithm; where* $\mathbf{K}, \widehat{\mathbf{K}}$ *are invertible upper triangular matrices and* $\mathbf{H}, \widehat{\mathbf{H}} \in \mathbb{SL}(3)\backslash\mathbb{SO}(3)$. *Then, the following condition*

$$\mathbf{w}(\mathbf{H}, \mathbf{K}, \mathbf{p}) = \mathbf{w}(\widehat{\mathbf{H}}, \widehat{\mathbf{K}}, \mathbf{p}),\ \forall \mathbf{p} \in \mathbb{R}^2, \tag{4.23}$$

is satisfied if and only if $\widehat{\mathbf{K}} = \mathbf{K}$ *and* $\widehat{\mathbf{H}} = \mathbf{H}$.

Proof of Theorem 1. Using (4.21) and (4.21) with $\xi = \widehat{\xi} = 1$ one obtains

$$\frac{\widehat{\mathbf{h}}_1^\top \widehat{\overline{\mathbf{X}}}_s}{\widehat{\mathbf{h}}_3^\top \widehat{\overline{\mathbf{X}}}_s - \|\widehat{\mathbf{H}}\widehat{\overline{\mathbf{X}}}_s\|} = m_{11}\frac{\mathbf{h}_1^\top \overline{\mathbf{X}}_s}{\mathbf{h}_3^\top \overline{\mathbf{X}}_s - \|\mathbf{H}\overline{\mathbf{X}}_s\|} + m_{12}\frac{\mathbf{h}_2^\top \overline{\mathbf{X}}_s}{\mathbf{h}_3^\top \overline{\mathbf{X}}_s - \|\mathbf{H}\overline{\mathbf{X}}_s\|} + m_{13} \tag{4.24}$$

$$\frac{\widehat{\mathbf{h}}_2^\top \widehat{\overline{\mathbf{X}}}_s}{\widehat{\mathbf{h}}_3^\top \widehat{\overline{\mathbf{X}}}_s - \|\widehat{\mathbf{H}}\widehat{\overline{\mathbf{X}}}_s\|} = m_{22}\frac{\mathbf{h}_2^\top \overline{\mathbf{X}}_s}{\mathbf{h}_3^\top \overline{\mathbf{X}}_s - \|\mathbf{H}\overline{\mathbf{X}}_s\|} + m_{23} \tag{4.25}$$

From (4.24) and (4.25) one verifies

$$\begin{aligned}\left(m_{11}\mathbf{h}_1^\top \overline{\mathbf{X}}_s + m_{12}\mathbf{h}_2^\top \overline{\mathbf{X}}_s + m_{13}\left(\mathbf{h}_3^\top \overline{\mathbf{X}}_s - \|\mathbf{H}\overline{\mathbf{X}}_s\|\right)\right)\left(\widehat{\mathbf{h}}_2^\top \widehat{\overline{\mathbf{X}}}_s\right)\\ = \left(m_{22}\mathbf{h}_2^\top \overline{\mathbf{X}}_s + m_{23}\left(\mathbf{h}_3^\top \overline{\mathbf{X}}_s - \|\mathbf{H}\overline{\mathbf{X}}_s\|\right)\right)\left(\widehat{\mathbf{h}}_1^\top \widehat{\overline{\mathbf{X}}}_s\right)\end{aligned} \tag{4.26}$$

As the reader can notice, we have to solve a non-linear system of $2N$ equations (N, number of pixels) with 14 unknown terms: 9 estimated homography parameters and 5 estimated intrinsic parameters (i.e. $\widehat{\mathbf{H}}$, $\mathbf{M}$). The objective is to prove that $\widehat{\mathbf{H}} = \mathbf{H}$ and $\mathbf{M} = \mathbf{I}$. Due to the square root terms $\|\mathbf{H}\overline{\mathbf{X}}_s\|$ and $\|\widehat{\mathbf{H}}\widehat{\overline{\mathbf{X}}}_s\|$ the problem is not trivial. Let us define the following bivariate polynomials of x, y for $i = 1, 2, 3$

$$\begin{cases} P_1^i \triangleq \mathbf{h}_i^\top \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \\ P_2^i \triangleq \mathbf{h}_i^\top \begin{pmatrix} x \\ y \\ -x^2 - y^2 \end{pmatrix} \\ \widehat{P}_1^i \triangleq \widehat{\mathbf{h}}_i^\top \begin{pmatrix} m_{11}x + m_{12}y + m_{13} \\ m_{22}y + m_{23} \\ 1 \end{pmatrix} \\ \widehat{P}_2^i \triangleq \widehat{\mathbf{h}}_i^\top \begin{pmatrix} m_{11}x + m_{12}y + m_{13} \\ m_{22}y + m_{23} \\ -(m_{11}x + m_{12}y + m_{13})^2 - (m_{22}y + m_{23})^2 \end{pmatrix} \end{cases} \tag{4.27}$$

In view of (4.26) and (4.27) one deduces the compact form of (4.26)

$$C_{2,1} + C_{3,1} + C_{3,2} + C_4 = (C_1 + C_{2,2})\sqrt{B_3 + B_4} \tag{4.28}$$

with

$$\begin{cases} B_3 \triangleq 2\sum_{i=1}^{3} P_1^i P_2^i \\ B_4 \triangleq \sum_{i=1}^{3}(P_1^i)^2 + \sum_{i=1}^{3}(P_2^i)^2 \end{cases} \tag{4.29}$$

and

$$\begin{cases} C_{2,1} & \triangleq & \widehat{P}_1^1 \sum_{i=2}^{3} m_{2i} P_1^i - \widehat{P}_1^2 \sum_{i=1}^{3} m_{1i} P_1^i \\ C_{3,1} & \triangleq & \widehat{P}_1^1 \sum_{i=2}^{3} m_{2i} P_2^i - \widehat{P}_1^2 \sum_{i=1}^{3} m_{1i} P_2^i \\ C_{3,2} & \triangleq & \widehat{P}_2^1 \sum_{i=2}^{3} m_{2i} P_1^i - \widehat{P}_2^2 \sum_{i=1}^{3} m_{1i} P_1^i \\ C_4 & \triangleq & \widehat{P}_2^1 \sum_{i=2}^{3} m_{2i} P_2^i - \widehat{P}_2^2 \sum_{i=1}^{3} m_{1i} P_2^i \\ C_1 & \triangleq & -m_{13}\widehat{P}_1^2 + m_{23}\widehat{P}_1^1 \\ C_{2,2} & \triangleq & -m_{13}\widehat{P}_2^2 + m_{23}\widehat{P}_2^1 \end{cases} \tag{4.30}$$

Equations (4.24), (4.25) and (4.28) will be used hereafter to deduce the unknown terms.

For the sake of the analysis, let us introduce the following properties whose proofs are given in the Appendix B. The proof of Property 3 is not given because is trivial.

Property 1 *Let A_2, B_2 be bivariate polynomials of order 2 and A_4, B_4 be bivariate polynomials of order 4. If A_2, B_2 are irreducible, and $\sqrt{A_4}$, $\sqrt{B_4}$ are not polynomials, and $A_2\sqrt{A_4} = B_2\sqrt{B_4}$, then there exists a non-null constant $\eta \in \mathbb{R}$ such that $A_2 = \eta B_2$ and $\sqrt{B_4} = \eta\sqrt{A_4}$.*

Property 2 *Let P_m , P_n and P_l be bivariate polynomials of order m , n and l, respectively, with $m, n, l \in \mathbb{N}^+$. If $P_m = P_n\sqrt{P_l}$ and $\sqrt{P_l}$ is not a bivariate polynomial, then P_m and P_n are identically null.*

Property 3 *Let P_m and P_n be two bivariate polynomials of order $m, n \in \mathbb{N}^+$. Then, $P_m P_n$ is identically null if and only if P_m or P_n is identically null.*

Property 4 *Any polynomial A_2 of the form: $a_1x + a_2y + a_3 + a_4(x^2 + y^2)$, with some constants $a_{1,2,3,4} \in \mathbb{R}$, is irreducible.*

In what follows, we calculate the solution of the unknown terms of the equations (4.24) and (4.25). Let us consider (4.28) and state Lemma 1. The proof of Lemma 1 is given in the Appendix B.

Lemma 1 *The term $\sqrt{B_3 + B_4}$ in (4.28) is a polynomial if and only if $\mathbf{H} \in \mathbb{SO}(3)$.*

From Lemma 1 and the assumption that $\mathbf{H} \notin \mathbb{SO}(3)$ made in Theorem 1, one deduces that $\sqrt{B_3 + B_4}$ is not a polynomial. This fact implies that $\sqrt{B_3 + B_4}$ is a polynomial only if the camera's movement is a *pure rotation*. If $\mathbf{H} \in \mathbb{SO}(3)$

then $\mathbf{H} \neq \mathbf{R} + \mathbf{t}\mathbf{n}^\top$. Therefore, if the camera's movement is a rotation combined with a translation then $\sqrt{B_3 + B_4}$ is not a polynomial. In [MV07], is shown how the two valid physically solutions of the homography decomposition are related. Using the notation in [MV07], one has $\mathbf{H} = \mathbf{R}_a + \mathbf{t}_a\mathbf{n}_a^\top = \mathbf{R}_b + \mathbf{t}_a\mathbf{n}_b^\top$ and $\mathbf{t}_b = \frac{\|\mathbf{t}_a\|}{\rho}\mathbf{R}_a(2\mathbf{n}_a + \mathbf{R}_a^\top\mathbf{t}_a)$. So, if $\mathbf{t}_a = \mathbf{0}$ then $\mathbf{t}_b = \mathbf{0}$. If $\mathbf{t}_b = \mathbf{0}$ then $\mathbf{t}_a = \mathbf{0}$ because $2\mathbf{n}_a \neq -\mathbf{R}_a^\top\mathbf{t}_a$ (see pages 37 to 38 in [MV07]). Nevertheless, since our calibration method and most other camera calibration methods perform calibration when the camera's movements are composed of a rotation combined with a translation, we will show the uniqueness of the solution for this kind of motion.

4.3.3.1 Camera's rotational coupled with translation movement

Since $\sqrt{B_4 + B_3}$ is not a polynomial, from (4.28) and Property 2, one has

$$\begin{cases} C_1 + C_{2,2} = 0 \\ C_{2,1} + C_{3,1} + C_{3,2} + C_4 = 0 \end{cases}$$

From $C_1 + C_{2,2} = 0$ and (4.30) one has

$$\begin{aligned} & m_{13}\widehat{\mathbf{h}}_2^\top \begin{pmatrix} 2(m_{11}x + m_{12}y + m_{13}) \\ 2(m_{22}y + m_{23}) \\ 1 - (m_{11}x + m_{12}y + m_{13})^2 - (m_{22}y + m_{23})^2 \end{pmatrix} \\ & = m_{23}\widehat{\mathbf{h}}_1^\top \begin{pmatrix} 2(m_{11}x + m_{12}y + m_{13}) \\ 2(m_{22}y + m_{23}) \\ 1 - (m_{11}x + m_{12}y + m_{13})^2 - (m_{22}y + m_{23})^2 \end{pmatrix} \end{aligned} \tag{4.31}$$

Equalizing the coefficients of x^2 and y^2 in the left-hand side (LHS) and right-hand side (RHS) of (4.31), one verifies that

$$\begin{cases} m_{13}m_{11}^2\widehat{h}_{23} = m_{23}m_{11}^2\widehat{h}_{13} \\ m_{13}(m_{12}^2 + m_{22}^2)\widehat{h}_{23} = m_{23}(m_{12}^2 + m_{22}^2)\widehat{h}_{13} \end{cases} \tag{4.32}$$

Then, from (4.32) and the fact that $m_{11} > 0$ one deduces

$$m_{13}\widehat{h}_{23} = m_{23}\widehat{h}_{13} \tag{4.33}$$

From (4.31) and (4.33) one has

$$m_{13}\begin{pmatrix}\widehat{h}_{21} \\ \widehat{h}_{22}\end{pmatrix}^\top \begin{pmatrix} m_{11}x + m_{12}y + m_{13} \\ m_{22}y + m_{23}\end{pmatrix} = m_{23}\begin{pmatrix}\widehat{h}_{11} \\ \widehat{h}_{12}\end{pmatrix}^\top \begin{pmatrix} m_{11}x + m_{12}y + m_{13} \\ m_{22}y + m_{23}\end{pmatrix} \tag{4.34}$$

Equalizing the coefficients of x, y and constant term in the LHS and RHS of (4.34), one obtains

$$m_{13}\widehat{h}_{21} = m_{23}\widehat{h}_{11} \tag{4.35}$$

$$m_{13}(m_{12}\widehat{h}_{21} + m_{22}\widehat{h}_{22}) = m_{23}(m_{12}\widehat{h}_{11} + m_{22}\widehat{h}_{12}) \tag{4.36}$$

$$m_{13}(m_{13}\widehat{h}_{21} + m_{23}\widehat{h}_{22}) = m_{23}(m_{13}\widehat{h}_{11} + m_{23}\widehat{h}_{12}) \tag{4.37}$$

From (4.35) , (4.36) and the fact that $m_{22} > 0$ one deduces

$$m_{13}\widehat{h}_{22} = m_{23}\widehat{h}_{12} \tag{4.38}$$

From Equations (4.33), (4.35) and (4.38) one obtains

$$m_{23}\begin{pmatrix}\widehat{h}_{11}\\ \widehat{h}_{12}\\ \widehat{h}_{13}\end{pmatrix} = m_{13}\begin{pmatrix}\widehat{h}_{21}\\ \widehat{h}_{22}\\ \widehat{h}_{23}\end{pmatrix} \tag{4.39}$$

From (4.39) one deduces that if either m_{13} or m_{23} is not null then $\widehat{\mathbf{h}}_1 \times \widehat{\mathbf{h}}_2 = 0$ which implies $\det(\widehat{\mathbf{H}}) = \widehat{\mathbf{h}}_3^\top(\widehat{\mathbf{h}}_1 \times \widehat{\mathbf{h}}_2) = 0$. But, this contradicts with the fact that $\det(\widehat{\mathbf{H}}) = 1$. As a consequence, one ensures that

$$m_{13} = m_{23} = 0 \tag{4.40}$$

Using the fact that $\sqrt{B_3 + B_4}$ $(= \|\mathbf{H}\overline{\mathbf{X}}_s\|$) is not a polynomial, we will prove that $\|\widehat{\mathbf{H}}\widehat{\overline{\mathbf{X}}}_s\|$ is not a polynomial neither. Let us prove this by contradiction. Assume that $\|\widehat{\mathbf{H}}\widehat{\overline{\mathbf{X}}}_s\|$ is a polynomial. Then, from (4.25) (with $m_{23} = 0$) one has

$$\widehat{\mathbf{h}}_2^\top\widehat{\overline{\mathbf{X}}}_s\left(\mathbf{h}_3^\top\overline{\mathbf{X}}_s - \|\mathbf{H}\overline{\mathbf{X}}_s\|\right) = m_{22}\mathbf{h}_2^\top\overline{\mathbf{X}}_s(\widehat{\mathbf{h}}_3^\top\widehat{\overline{\mathbf{X}}}_s - \|\widehat{\mathbf{H}}\widehat{\overline{\mathbf{X}}}_s\|)$$

which is equivalent to

$$\widehat{\mathbf{h}}_2^\top\widehat{\overline{\mathbf{X}}}_s\mathbf{h}_3^\top\overline{\mathbf{X}}_s - m_{22}\mathbf{h}_2^\top\overline{\mathbf{X}}_s(\widehat{\mathbf{h}}_3^\top\widehat{\overline{\mathbf{X}}}_s - \|\widehat{\mathbf{H}}\widehat{\overline{\mathbf{X}}}_s\|) = \widehat{\mathbf{h}}_2^\top\widehat{\overline{\mathbf{X}}}_s\|\mathbf{H}\overline{\mathbf{X}}_s\| \tag{4.41}$$

From (4.41), the facts that $\|\widehat{\mathbf{H}}\widehat{\overline{\mathbf{X}}}_s\|$ is a polynomial and $\|\mathbf{H}\overline{\mathbf{X}}_s\|$ is not a polynomial, and Property 2, one deduces that $\widehat{\mathbf{h}}_2^\top\widehat{\overline{\mathbf{X}}}_s \equiv 0$. This is impossible since this equality is not verified for all x, y. This can be directly verified using (4.13) and the fact that $\widehat{\mathbf{h}}_2 \neq 0$. The resulting contradiction implies that $\|\widehat{\mathbf{H}}\widehat{\overline{\mathbf{X}}}_s\|$ is not a polynomial. From here, let us continue to study (4.41). First, (4.41) can be rewritten in a more compact form as

$$D_4 + A_{2,1}\sqrt{A_{4,1}} = A_{2,2}\sqrt{A_{4,2}}, \tag{4.42}$$

with $A_{2,1} \triangleq m_{22}\mathbf{h}_2^\top\overline{\mathbf{X}}_s$, $A_{2,2} \triangleq \widehat{\mathbf{h}}_2^\top\widehat{\overline{\mathbf{X}}}_s$, $A_{4,1} \triangleq \|\widehat{\mathbf{H}}\widehat{\overline{\mathbf{X}}}_s\|^2$, $A_{4,2} \triangleq \|\mathbf{H}\overline{\mathbf{X}}_s\|^2$, and

$$D_4 \triangleq \widehat{\mathbf{h}}_2^\top\widehat{\overline{\mathbf{X}}}_s\mathbf{h}_3^\top\overline{\mathbf{X}}_s - m_{22}\mathbf{h}_2^\top\overline{\mathbf{X}}_s\widehat{\mathbf{h}}_3^\top\widehat{\overline{\mathbf{X}}}_s.$$

One easily verifies that $A_{2,1}$ and $A_{2,2}$ are not identically null. Powering the LHS and RHS of (4.42) by two, one deduces

$$D_4^2 + A_{2,1}^2 A_{4,1} - A_{2,2}^2 A_{4,2} = -2D_4 A_{2,1}\sqrt{A_{4,1}}.$$

Using this relation and the fact that $\sqrt{A_{4,1}}$ is not a polynomial, one verifies from Property 2 that $D_4 A_{2,1} \equiv 0$. Then, using the latter relation and the fact that $A_{2,1}$ is not identically null, one verifies from Property 3 that $D_4 \equiv 0$ which in turn implies that

$$A_{2,1}\sqrt{A_{4,1}} = A_{2,2}\sqrt{A_{4,2}}. \tag{4.43}$$

From the definitions of $\overline{\mathbf{X}}_s$ and $\widehat{\overline{\mathbf{X}}}_s$ and from Property 4, one can verify that all polynomials $\mathbf{h}_i^\top \overline{\mathbf{X}}_s$ and $\widehat{\mathbf{h}}_i^\top \widehat{\overline{\mathbf{X}}}_s$, with $i \in \{1,2,3\}$ are irreducible. This indicates that $A_{2,1}$ and $A_{2,2}$ are irreducible and non null (proved previously). Using this and the facts that $\sqrt{A_{4,1}}$ and $\sqrt{A_{4,2}}$ are not polynomials, one verifies from Property 1 and (4.43) the existence of a constant $\eta \neq 0$ such that

$$\|\widehat{\mathbf{H}}\widehat{\overline{\mathbf{X}}}_s\| = \eta \|\mathbf{H}\overline{\mathbf{X}}_s\| \tag{4.44}$$

Then, from (4.41) and Property 2, one deduces that

$$(\widehat{\mathbf{h}}_2^\top \widehat{\overline{\mathbf{X}}}_s)(\mathbf{h}_3^\top \overline{\mathbf{X}}_s) = m_{22}(\mathbf{h}_2^\top \overline{\mathbf{X}}_s)(\widehat{\mathbf{h}}_3^\top \widehat{\overline{\mathbf{X}}}_s) \tag{4.45}$$

$$(\widehat{\mathbf{h}}_2^\top \widehat{\overline{\mathbf{X}}}_s) = \eta\, m_{22}(\mathbf{h}_2^\top \overline{\mathbf{X}}_s) \tag{4.46}$$

Replacing (4.40) into (4.46), one obtains

$$\begin{pmatrix} \widehat{h}_{21} \\ \widehat{h}_{22} \\ \widehat{h}_{23} \end{pmatrix}^\top \begin{pmatrix} 2(m_{11}x + m_{12}y) \\ 2m_{22}y \\ 1 - (m_{11}x + m_{12}y)^2 - m_{22}^2 y^2 \end{pmatrix} = \eta m_{22} \begin{pmatrix} h_{21} \\ h_{22} \\ h_{23} \end{pmatrix}^\top \begin{pmatrix} 2x \\ 2y \\ 1 - x^2 - y^2 \end{pmatrix} \tag{4.47}$$

Equalizing the constant terms in the LHS and RHS of (4.47), one gets

$$\widehat{h}_{23} = \eta m_{22} h_{23} \tag{4.48}$$

For all circumstances, there exist only two possible cases:

- **Case 1:** $h_{23} \neq 0$.
- **Case 2:** $h_{23} = 0$.

Now, let us consider the Case 1 where $h_{23} \neq 0$:

From (4.48), if $h_{23} \neq 0$ then $\widehat{h}_{23} \neq 0$ since $\eta \neq 0$ and $m_{22} > 0$. Then, equalizing the coefficients of xy, x^2, and y^2 in the LHS and RHS of (4.47), one obtains

$$\begin{cases} m_{11}m_{12} = 0 \\ m_{11}^2 = 1 \\ m_{12}^2 + m_{22}^2 = 1 \end{cases} \tag{4.49}$$

From (4.49) and the facts that $m_{11} > 0$ and $m_{22} > 0$, one easily deduces

$$\begin{cases} m_{11} = m_{22} = 1 \\ m_{12} = 0 \end{cases} \tag{4.50}$$

From (4.40) and (4.50), one already obtains $\mathbf{M} = \mathbf{I}$ which in turn indicates that $\widehat{\mathbf{K}} = \mathbf{K}$. It remains to prove that $\widehat{\mathbf{H}} = \mathbf{H}$. From (4.47) and (4.50), one deduces

$$\begin{cases} \widehat{\overline{\mathbf{X}}}_s = \overline{\mathbf{X}}_s \\ \widehat{\mathbf{h}}_2 = \eta \mathbf{h}_2 \end{cases} \tag{4.51}$$

From (4.45) and (4.51), one deduces $\eta \mathbf{h}_3^\top \overline{\mathbf{X}}_s = \widehat{\mathbf{h}}_3^\top \widehat{\overline{\mathbf{X}}}_s = \widehat{\mathbf{h}}_3^\top \overline{\mathbf{X}}_s$, which implies that

$$\widehat{\mathbf{h}}_3 = \eta \mathbf{h}_3.$$

Replacing (4.40), (4.50), (4.44), and (4.51) into (4.24), one obtains

$$\frac{\widehat{\mathbf{h}}_1^\top \overline{\mathbf{X}}_s}{\eta(\mathbf{h}_3^\top \overline{\mathbf{X}}_s - \|\mathbf{H}\overline{\mathbf{X}}_s\|)} = \frac{\mathbf{h}_1^\top \overline{\mathbf{X}}_s}{\mathbf{h}_3^\top \overline{\mathbf{X}}_s - \|\mathbf{H}\overline{\mathbf{X}}_s\|}.$$

From here, using Property 2 and the fact that $\|\mathbf{H}\overline{\mathbf{X}}_s\|$ is not a polynomial, one deduces $\widehat{\mathbf{h}}_1^\top \overline{\mathbf{X}}_s = \eta \mathbf{h}_1^\top \overline{\mathbf{X}}_s$, which in turn yields $\widehat{\mathbf{h}}_1 = \eta \mathbf{h}_1$. Then, using the fact that $\det(\mathbf{H}) = \det(\widehat{\mathbf{H}}) = 1$, one deduces

$$\widehat{\mathbf{h}}_1^\top (\widehat{\mathbf{h}}_2 \times \widehat{\mathbf{h}}_3) = \eta \mathbf{h}_1^\top (\eta \mathbf{h}_2 \times \eta \mathbf{h}_3) = 1.$$

Therefore, one can deduce $\eta^3 = 1$ and, subsequently, $\eta = 1$. This implies that $\widehat{\mathbf{h}}_1 = \mathbf{h}_1$, $\widehat{\mathbf{h}}_2 = \mathbf{h}_2$, and $\widehat{\mathbf{h}}_3 = \mathbf{h}_3$, or equivalently $\widehat{\mathbf{H}} = \mathbf{H}$.

Now, let us consider Case 2 where $h_{23} = 0$:

From (4.48), one deduces that $\widehat{h}_{23} = 0$ since $\eta \neq 0$. Then, using (4.40) and (4.44) and equalizing the coefficients of x and y in the LHS and RHS of (4.47) one deduces

$$m_{11}\widehat{h}_{21} = \eta m_{22} h_{21}, \tag{4.52}$$

$$m_{12}\widehat{h}_{21} + m_{22}\widehat{h}_{22} = \eta m_{22} h_{22}. \tag{4.53}$$

From (4.24), (4.40), (4.44) and Property 2, one deduces

$$(\widehat{\mathbf{h}}_1^\top \widehat{\overline{\mathbf{X}}}_s) = \eta(m_{11}\mathbf{h}_1^\top \overline{\mathbf{X}}_s + m_{12}\mathbf{h}_2^\top \overline{\mathbf{X}}_s), \tag{4.54}$$

$$(\widehat{\mathbf{h}}_2^\top \widehat{\overline{\mathbf{X}}}_s)(\mathbf{h}_3^\top \overline{\mathbf{X}}_s) = (m_{11}\mathbf{h}_1^\top \overline{\mathbf{X}}_s + m_{12}\mathbf{h}_2^\top \overline{\mathbf{X}}_s)(\widehat{\mathbf{h}}_3^\top \widehat{\overline{\mathbf{X}}}_s). \tag{4.55}$$

From (4.54) one has

$$\begin{pmatrix} \widehat{h}_{11} \\ \widehat{h}_{12} \\ \widehat{h}_{13} \end{pmatrix}^\top \begin{pmatrix} 2(m_{11}x + m_{12}y) \\ 2m_{22}y \\ 1 - (m_{11}x + m_{12}y)^2 - m_{22}^2 y^2 \end{pmatrix} = \eta \begin{pmatrix} m_{11}h_{11} + m_{12}h_{21} \\ m_{11}h_{12} + m_{12}h_{22} \\ m_{11}h_{13} + m_{12}h_{23} \end{pmatrix}^\top \begin{pmatrix} 2x \\ 2y \\ 1 - x^2 - y^2 \end{pmatrix} \tag{4.56}$$

Equalizing the constant terms in the LHS and RHS of (4.56), one obtains

$$\widehat{h}_{13} = \eta m_{11} h_{13}, \tag{4.57}$$

which conveys to the two following subcases:

- Subcase 2.1: $h_{13} \neq 0$.
- Subcase 2.2: $h_{13} = 0$.

Consider Subcase 2.1 where $h_{13} \neq 0$:

From (4.57), if $h_{13} \neq 0$ then $\widehat{h}_{13} \neq 0$ since $\eta, m_{11} \neq 0$. Therefore, from the coefficient of xy, x^2 and y^2 of (4.56) one deduces

$$\widehat{h}_{13} m_{11} m_{12} = 0, \tag{4.58}$$

$$m_{11}^2 \widehat{h}_{13} = \eta(m_{11} h_{13} + m_{12} h_{23}), \tag{4.59}$$

$$(m_{22}^2 + m_{12}^2)\widehat{h}_{13} = \eta(m_{11} h_{13} + m_{12} h_{23}). \tag{4.60}$$

From (4.57), (4.58), (4.59), (4.60) and the fact that $\widehat{h}_{13} \neq 0$, $m_{11} > 0$ and $m_{22} > 0$ one deduces

$$\begin{cases} m_{11} = m_{22} = 1 \\ m_{12} = 0 \end{cases} \tag{4.61}$$

Subsequently, from (4.40) and (4.61) one deduces $\mathbf{M} = \mathbf{I}$, so that $\widehat{\mathbf{K}} = \mathbf{K}$. From (4.61) and (4.56), one deduces

$$\begin{cases} \widehat{\overline{\mathbf{X}}}_s = \overline{\mathbf{X}}_s \\ \widehat{\mathbf{h}}_1 = \eta \mathbf{h}_1 \end{cases} \tag{4.62}$$

From (4.52), (4.53), and (4.54), one deduces $\widehat{\mathbf{h}}_2 = \eta \mathbf{h}_2$ and $\widehat{\mathbf{h}}_3 = \eta \mathbf{h}_3$. From here, analogously to Case 1, one can deduce that $\eta = 1$ and, subsequently, $\widehat{\mathbf{H}} = \mathbf{H}$.

Finally, consider Subcase 2.2 where $h_{13} = 0$:

From (4.57) if $h_{13} = 0$ then $\widehat{h}_{13} = 0$. Then, from (4.56) and the fact that $h_{23} = 0$, one has

$$\begin{pmatrix} \widehat{h}_{11} \\ \widehat{h}_{12} \end{pmatrix}^{\top} \begin{pmatrix} m_{11}x + m_{12}y \\ m_{22}y \end{pmatrix} = \eta \begin{pmatrix} m_{11}h_{11} + m_{12}h_{21} \\ m_{11}h_{12} + m_{12}h_{22} \end{pmatrix}^{\top} \begin{pmatrix} x \\ y \end{pmatrix} \tag{4.63}$$

From the coefficients of x and y in (4.57) and (4.63) one has

$$m_{11} \widehat{h}_{21} = \eta m_{22} h_{21}, \tag{4.64}$$

$$m_{12} \widehat{h}_{21} + m_{22} \widehat{h}_{22} = \eta m_{22} h_{22}, \tag{4.65}$$

$$m_{11} \widehat{h}_{11} = \eta(m_{11} h_{11} + m_{12} h_{21}), \tag{4.66}$$

$$m_{12}\widehat{h}_{11} + m_{22}\widehat{h}_{12} = \eta(m_{11}h_{12} + m_{12}h_{22}). \tag{4.67}$$

From the fact $\widehat{h}_{13} = h_{13} = 0$, $\det(\mathbf{H}) = \det(\widehat{\mathbf{H}}) = 1$, $m_{11} > 0$, $m_{22} > 0$ and equations (4.44), (4.64), (4.65), (4.66), (4.67) it is straightforward to find that $\mathbf{M} = \mathbf{I}$ and $\eta = 1$, and, subsequently $\widehat{\mathbf{K}} = \mathbf{K}, \widehat{\mathbf{H}} = \mathbf{H}$ *(end of proof of Theorem 1).*

Remark: The proof of Theorem 1 ensures the uniqueness of solution of the estimated parameters for the calibration method proposed in this thesis when the camera's movement is composed of a rotation combined with a non-null translation. The technique presented here could be extended for methods involving similar calibration parameters and projection model [MBMR06, WLH06, DWW07, YZ05, BA05, Vas04, YH03].

4.4 Conclusion

We have shown a direct approach to the self-calibration of omnidirectional cameras. Synthetic data have shown that under certainly conditions, no illumination changes and blur effect, the estimated parameters converge to the correct value after a few images. On other hand, experiment with real data have shown that the estimation of parameters can be unobservables because of noise.

We have studied the uniqueness of the solution for the calibration problem of central catadioptric omnidirectional cameras when a planar mirror or a parabolic mirror are used. We have showed that when a planar mirror is used the solution could not be unique. In the case when parabolic mirror is used we have showed that the solution is unique when the camera's movement is a rotation plus a translation.

Chapter 5
Conclusion and Future research

Omnidirectional cameras have provided important improvements in recent research in computer vision and mobile robotics. The field of view of omnidirectional cameras is well suited for motion estimation and obstacle avoidance. Objects do not dissapear from the field of view but only change their image positions. Therefore, it is much easier to deal with applications such as selection of landmarks, tracking and motion detection. However, the use of an omnidirectional camera has always been limited by the calibration process. Most of the approaches to calibrate catadioptric omnidirectional cameras require awkward calibration steps or need known 3D calibration patterns that unable an on-line self-calibration.

This thesis addressed some of the issues of calibrating catadioptric omndirectional cameras by a novel approach: tracking of a plane in an image sequence. The image sequence can be the one used by the robot to performe its task. The choice for considering planes for calibrating cameras is because planes are very common in man-made environments, and often easily identifiable and rather accurately planar. Furthermore, planes are simple to process and allow very reliable and precise intensity-based matching, by fitting the homographies between image pairs.

The unified projection model by Geyer and Barreto was adapted to enable calibration of a wide variety of sensors used in robotics and computer vision. Compared to other models present in the literature, this model has easily identifiable parameters and presents a compromise between generecity and over-parameterization.

In this thesis, we neglected radial and decentering distortion because we observed that a complex projection model makes the cost function in the optimazation problem highly non-linear. Therefore, it not only would not help us but also would cause numerical instability.

This thesis also analysed visual tracking using an uncalibrated omnidirectional camera. We showed how visual tracking could be adapted to uncalibrated central catadioptric omndirectional cameras by using parametric models (in this case homographies) combined with the spherical perspective model. The approach can be used in applications where camera calibration is not necessary.

The convergence to the correct value of the estimated parameters is affected by illunination changes and blur effects. Therefore, it is recomended that the user provides image sequences with neither strong illumination changes nor blur effects.

We formulated the problem of uniqueness of the solution in the framework of a

non-linear equation system, such that the solution of this equations confirms the uniqueness of the solution. The framework we formulated was for the general case (all catadioptric sensors). We solved these general equations under certain conditions, namely for a paraboloid mirror. So, we were able to show analytically that the solution is unique when the camera's movement is composed of a rotation plus coupled with a translation. We remark here that in the greatest of cases robotics applications tend to use paraboloid mirrors, therefore this proof is of great importance. However, for the most general case the equations are complex and very hard to solve. In fact we were not capable of solving the general equations without the particular assumption on the mirror and a specific motion, i.e. the case we have shown. Therefore, the general case still remains unsolved. We suspect, like most in our community, that it is true even in the general case, but even in our formulation, the uniqueness of the solution still remains an unsolved problem.

5.1 Future Research

In Chapter 3, we discussed an efficient way of tracking a planar shape with an uncalibrated camera. SSD is however limited to small inter-frame motion. Ways of initialising the tracking, extending the convergence domain and recovering from total occlusion are important steps for applying the approach in general real-world situations.

In Chapter 4, we were not able to explore in detail the problem of illumination changes. Robust techniques such as Tukey or Huber were considered in order to improve the quality of the estimates. However, the results did not improve. On the contrary, the number of iterations increased drastically. Techniques such the one proposed by Silveira and Malis can be considered to improve the quality of the estimates and reduce the number of iterations.

The proposed calibration algorithm does not include lens distortions -radial and tangential. It would be useful to have a self-calibration method that could estimate these parameters.

In Chapter 4, we studied the uniqueness of the solution for calibrating catadioptric omnidirectional cameras. In spite of the efforts to prove for all mirror shapes, we have not achieved the task. Thus, a harder mathematical work can be considered with the objective of provides the solution for the general case.

Appendix A

Current Jacobian and Reference Jacobian

A.1 Current Jacobian

The current Jacobian is the derivative of the cost function in (3.4) when $\mathbf{x} = (\mathbf{z}, \boldsymbol{\xi}, \boldsymbol{\gamma}) = \mathbf{0}$. Therefore we must computate the following derivative

$$\mathbf{J}(\mathbf{0}) = \left[\nabla_{\mathbf{x}}\mathbf{I}'\left(\mathbf{w}\left(\widehat{\mathbf{H}}\mathbf{H}(\mathbf{z}), \widehat{\boldsymbol{\xi}} + \boldsymbol{\xi}, \widehat{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{p_i}\right)\right) - \mathbf{I}(\mathbf{p})\right]_{\mathbf{x}=\mathbf{0}}$$

The function I' can be expanded as follows

$$I'\left(\mathbf{w}\left(\widehat{\mathbf{H}}\mathbf{H}(\mathbf{z}), \widehat{\boldsymbol{\xi}} + \boldsymbol{\xi}, \widehat{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{p}\right)\right) = I'(\mathbf{w}(\widehat{\mathbf{H}}, \widehat{\boldsymbol{\xi}}, \widehat{\boldsymbol{\gamma}}, \mathbf{w}^{-1}(\widehat{\mathbf{H}}, \widehat{\boldsymbol{\xi}}, \widehat{\boldsymbol{\gamma}}, \mathbf{w}(\widehat{\mathbf{H}}\mathbf{H}(\mathbf{z}), \widehat{\boldsymbol{\xi}}+\boldsymbol{\xi}, \widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma}, \mathbf{p}))))$$

Let $\mathbf{q} = \mathbf{w}^{-1}(\widehat{\mathbf{H}}, \widehat{\boldsymbol{\xi}}, \widehat{\boldsymbol{\gamma}}, \mathbf{w}(\widehat{\mathbf{H}}\mathbf{H}(\mathbf{z}), \widehat{\boldsymbol{\xi}} + \boldsymbol{\xi}, \widehat{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{p}))$. Thus, using the chain rule for derivatives one obtains

$$\left.\frac{\partial I'(\mathbf{w}(\widehat{\mathbf{H}}, \widehat{\xi}, \widehat{\boldsymbol{\gamma}}, \mathbf{q}))}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_0} = \left.\frac{\partial I'(\mathbf{w}(\widehat{\mathbf{H}}, \widehat{\xi}, \widehat{\boldsymbol{\gamma}}, \mathbf{q}))}{\partial \mathbf{q}}\right|_{\mathbf{q}=\mathbf{p}} \left.\frac{\partial \mathbf{w}^{-1}(\widehat{\mathbf{H}}, \widehat{\boldsymbol{\xi}}, \widehat{\boldsymbol{\gamma}}, \mathbf{w}(\widehat{\mathbf{H}}\mathbf{H}(\mathbf{z}), \widehat{\boldsymbol{\xi}} + \boldsymbol{\xi}, \widehat{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{p}))}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{0}}$$

It is important to note that the first part of the derivative is a (1×2) vector that can be computed directly from the current image data. The vector contains the gradient of the warped image (i.e. the image being warped with $\mathbf{w}(\widehat{\mathbf{H}}, \widehat{\boldsymbol{\xi}}, \widehat{\boldsymbol{\gamma}}, \mathbf{p})$ computed at $\mathbf{p}$). This derivative will be noted $\mathbf{J}_{I'}$. Let $\mathbf{r} = \mathbf{w}(\widehat{\mathbf{H}}\mathbf{H}(\mathbf{z}), \widehat{\boldsymbol{\xi}} + \boldsymbol{\xi}, \widehat{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{p})$, thus the second part of the derivative can be decomposed into two parts as follows

$$\left.\frac{\partial \mathbf{w}^{-1}(\widehat{\mathbf{H}}, \widehat{\boldsymbol{\xi}}, \widehat{\boldsymbol{\gamma}}, \mathbf{r})}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{x}_0} = \left.\frac{\partial \mathbf{w}^{-1}(\widehat{\mathbf{H}}, \widehat{\boldsymbol{\xi}}, \widehat{\boldsymbol{\gamma}}, \mathbf{r})}{\partial \mathbf{r}}\right|_{\mathbf{r}=\mathbf{p}'} \left.\frac{\partial \mathbf{w}(\widehat{\mathbf{H}}\mathbf{H}(\mathbf{z}), \widehat{\boldsymbol{\xi}} + \boldsymbol{\xi}, \widehat{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{p})}{\partial \mathbf{x}}\right|_{\mathbf{x}=\mathbf{0}} \tag{A.1}$$

The first part of Equation (A.1) is

$$\left.\frac{\partial \mathbf{w}^{-1}(\widehat{\mathbf{H}}, \widehat{\boldsymbol{\xi}}, \widehat{\boldsymbol{\gamma}}, \mathbf{r})}{\partial \mathbf{r}}\right|_{\mathbf{r}=\mathbf{p}'} = \left(\left.\frac{\partial \mathbf{w}(\widehat{\mathbf{H}}, \widehat{\boldsymbol{\xi}}, \widehat{\boldsymbol{\gamma}}, \mathbf{r})}{\partial \mathbf{r}}\right|_{\mathbf{r}=\mathbf{p}}\right)^{-1} \tag{A.2}$$

This derivative is a (2×2) matrix which represents the coordinates variation from a point $\mathbf{p}'$ in the current image I' with respect to the coordinates of a point $\mathbf{p}$ in the reference image I. This derivative will be noted $\mathbf{J_w}$. In order to compute this derivative, we will use the warping expression from (3.2), thus the derivative to be computed is

$$\left.\frac{\partial \mathbf{w}(\widehat{\mathbf{H}}, \widehat{\boldsymbol{\xi}}, \widehat{\boldsymbol{\gamma}}, \mathbf{c}(\widehat{\xi} + \xi, \widehat{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{s}(\widehat{\mathbf{H}}\mathbf{H}(\mathbf{z}), \mathbf{c}^{-1}(\widehat{\boldsymbol{\xi}} + \boldsymbol{\xi}, \widehat{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{p}))))}{\partial \mathbf{r}}\right|_{\mathbf{r}=\mathbf{p}}$$

Thus, using the chain rule for derivatives one obtains

$$\left.\frac{\partial \mathbf{w}(\widehat{\mathbf{H}}, \widehat{\boldsymbol{\xi}}, \widehat{\boldsymbol{\gamma}}, \mathbf{r})}{\partial \mathbf{r}}\right|_{\mathbf{r}=\mathbf{p}} = \left.\frac{\partial \mathbf{c}(\widehat{\boldsymbol{\xi}} + \boldsymbol{\xi}, \widehat{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{A})}{\partial \mathbf{A}}\right|_{\mathbf{A}=\mathbf{X}'_s} \left.\frac{\partial \mathbf{s}(\widehat{\mathbf{H}}\mathbf{H}(\mathbf{z}), \mathbf{B})}{\partial \mathbf{B}}\right|_{\mathbf{B}=\mathbf{X}_s} \left.\frac{\partial \mathbf{c}^{-1}(\widehat{\boldsymbol{\xi}} + \boldsymbol{\xi}, \widehat{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{q})}{\partial \mathbf{r}}\right|_{\mathbf{r}=\mathbf{p}}$$

For the purpose of saving space and to show the entire current Jacobian in a single row, we will denote these derivatives as $\mathbf{J_c}$, $\mathbf{J_s}$ and $\mathbf{J_{c^{-1}}}$. The first part is

$$\mathbf{J_c} = \left.\frac{\partial \mathbf{c}(\widehat{\boldsymbol{\xi}} + \boldsymbol{\xi}, \widehat{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{A})}{\partial \mathbf{A}}\right|_{\mathbf{A}=\mathbf{X}'_s} = \left.\frac{\partial \mathbf{k}(\widehat{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{h}(\widehat{\boldsymbol{\xi}} + \boldsymbol{\xi}, \mathbf{A}))}{\partial \mathbf{A}}\right|_{\mathbf{A}=\mathbf{X}'_s}$$

Then one obtains

$$\mathbf{J_c} = \left.\frac{\partial \mathbf{c}(\widehat{\boldsymbol{\xi}} + \boldsymbol{\xi}, \widehat{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{A})}{\partial \mathbf{A}}\right|_{\mathbf{A}=\mathbf{X}'_s} = \left.\frac{\partial \mathbf{k}(\widehat{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{a})}{\partial \mathbf{a}}\right|_{\mathbf{a}=\mathbf{q}'} \left.\frac{\partial \mathbf{h}(\widehat{\boldsymbol{\xi}} + \boldsymbol{\xi}, \mathbf{A})}{\partial \mathbf{A}}\right|_{\mathbf{A}=\mathbf{X}'_s}$$

The second part is

$$\mathbf{J_s} = \left.\frac{\partial \mathbf{s}(\widehat{\mathbf{H}}\mathbf{H}(\mathbf{z}), \mathbf{B})}{\partial \mathbf{B}}\right|_{\mathbf{B}=\mathbf{X}_s}$$

The third part is

$$\mathbf{J_{c^{-1}}} = \left.\frac{\partial \mathbf{c}^{-1}(\widehat{\boldsymbol{\xi}} + \boldsymbol{\xi}, \widehat{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{r})}{\partial \mathbf{r}}\right|_{\mathbf{r}=\mathbf{p}} = \left.\frac{\partial \mathbf{h}^{-1}(\widehat{\boldsymbol{\xi}} + \boldsymbol{\xi}, \mathbf{k}^{-1}(\widehat{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{r}))}{\partial \mathbf{r}}\right|_{\mathbf{r}=\mathbf{p}}$$

Then one obtains

$$\mathbf{J}_{\mathbf{c}^{-1}} = \left.\frac{\partial \mathbf{c}^{-1}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi}, \widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma}, \mathbf{r})}{\partial \mathbf{r}}\right|_{\mathbf{r}=\mathbf{p}} = \left.\frac{\partial \mathbf{h}^{-1}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi}, \mathbf{b})}{\partial \mathbf{b}}\right|_{\mathbf{b}=\mathbf{q}} \left.\frac{\partial \mathbf{k}^{-1}(\widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma}, \mathbf{r})}{\partial \mathbf{r}}\right|_{\mathbf{r}=\mathbf{p}}$$

This last derivative allows us to write Equation (A.2) as the inverse of the product of three Jacobians.

$$\mathbf{J}_{\mathbf{w}} = \left(\mathbf{J}_{\mathbf{c}}\mathbf{J}_{\mathbf{s}}\mathbf{J}_{\mathbf{c}^{-1}}\right)^{-1} \tag{A.3}$$

Now, we procede to compute the second part of Equation (A.1). This derivative will be computed considering the different sets of parameters in the warping function (mirror parameter, intrinsic parameters and homography parameters). This process allows to find a modular Jacobian which will be easily modifiable for the set or number of parameters to be estimated. Let us start with the homography parameters.

A.1.1 Homography parameters Jacobian

The derivative to be computed is

$$\left.\frac{\partial \mathbf{c}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi}, \widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma}, \mathbf{s}(\widehat{\mathbf{H}}\mathbf{H}(\mathbf{z}), \mathbf{c}^{-1}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi}, \widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma}, \mathbf{p})))}{\partial \mathbf{z}}\right|_{\mathbf{z}=\mathbf{z}_0}$$

Let $\mathbf{A} = \mathbf{s}(\widehat{\mathbf{H}}\mathbf{H}(\mathbf{z}), \mathbf{c}^{-1}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi}, \widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma}, \mathbf{p}))$. Thus one has

$$\left.\frac{\partial \mathbf{c}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi}, \widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma}, \mathbf{A}))}{\partial \mathbf{z}}\right|_{\mathbf{z}=\mathbf{z}_0} = \left.\frac{\partial \mathbf{c}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi}, \widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma}, \mathbf{A})}{\partial \mathbf{A}}\right|_{\mathbf{A}=\mathbf{X}_s'} \left.\frac{\partial \mathbf{s}(\widehat{\mathbf{H}}\mathbf{H}(\mathbf{z}), \mathbf{c}^{-1}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi}, \widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma}, \mathbf{p}))}{\partial \mathbf{z}}\right|_{\mathbf{z}=\mathbf{z}_0}$$

The first part is the Jacobian $\mathbf{J}_{\mathbf{c}}$ which has already been computed. The second part is

$$\left.\frac{\partial \mathbf{s}(\widehat{\mathbf{H}}\mathbf{H}(\mathbf{z}), \mathbf{c}^{-1}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi}, \widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma}, \mathbf{p}))}{\partial \mathbf{z}}\right|_{\mathbf{z}=\mathbf{z}_0} = \left.\frac{\partial \mathbf{s}(\widehat{\mathbf{H}}\mathbf{H}(\mathbf{z}), \mathbf{c}^{-1}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi}, \widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma}, \mathbf{p}))}{\partial \mathbf{H}(\mathbf{z})}\right|_{\mathbf{H}(\mathbf{z})=\mathbf{I}} \left.\frac{\partial \widehat{\mathbf{H}}\mathbf{H}(\mathbf{z})}{\partial \mathbf{z}}\right|_{\mathbf{z}=\mathbf{z}_0}$$

These derivatives end up with the homography parameters Jacobian. We will note them as $\mathbf{J}_{\mathbf{s}_\mathbf{H}}$ and $\mathbf{J}_{\mathbf{H}_\mathbf{z}}$. Thus, the homography parameters Jacobian can be written as

$$\mathbf{J}_{I'}\mathbf{J}_{\mathbf{w}}\mathbf{J}_{\mathbf{c}}\mathbf{J}_{\mathbf{s}_\mathbf{H}}\mathbf{J}_{\mathbf{H}_\mathbf{z}}(\mathbf{0}) \tag{A.4}$$

A.1.2 Mirror parameter Jacobian

Now, we will compute the mirror parameter Jacobian. Similarly to the last Jacobian, the derivative to be computed is

$$\left.\frac{\partial \mathbf{c}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi},\widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma},\mathbf{s}(\widehat{\mathbf{H}}\mathbf{H}(\mathbf{z}),\mathbf{c}^{-1}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi},\widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma},\mathbf{p})))}{\partial \boldsymbol{\xi}}\right|_{\boldsymbol{\xi}=\boldsymbol{\xi}_0}$$

Again, let $\mathbf{A}=\mathbf{s}(\widehat{\mathbf{H}}\mathbf{H}(\mathbf{z}),\mathbf{c}^{-1}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi},\widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma},\mathbf{p}))$, thus one has

$$\left.\frac{\partial \mathbf{c}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi},\widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma},\mathbf{X}_s))}{\partial \boldsymbol{\xi}}\right|_{\boldsymbol{\xi}=\boldsymbol{\xi}_0}=\left.\frac{\partial \mathbf{c}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi},\widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma},\mathbf{A})}{\partial \boldsymbol{\xi}}\right|_{\boldsymbol{\xi}=\boldsymbol{\xi}_0}+\mathbf{J}_{\mathbf{c}}\mathbf{J}_{\mathbf{s}}\left.\frac{\partial \mathbf{c}^{-1}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi},\widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma},\mathbf{p})}{\partial \boldsymbol{\xi}}\right|_{\boldsymbol{\xi}=\boldsymbol{\xi}_0}$$

The first part is

$$\mathbf{J}_{\mathbf{c}_{\boldsymbol{\xi}}}=\left.\frac{\partial \mathbf{c}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi},\widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma},\mathbf{X}_s)}{\partial \boldsymbol{\xi}}\right|_{\boldsymbol{\xi}=\boldsymbol{\xi}_0}=\left.\frac{\partial \mathbf{k}(\widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma},\mathbf{a})}{\partial \mathbf{a}}\right|_{\mathbf{a}=\mathbf{q}'}\left.\frac{\partial \mathbf{h}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi},\mathbf{X}_s)}{\partial \boldsymbol{\xi}}\right|_{\boldsymbol{\xi}=\boldsymbol{\xi}_0}$$

The last part is

$$\mathbf{J}_{\mathbf{c}_{\boldsymbol{\xi}}^{-1}}=\left.\frac{\partial \mathbf{c}^{-1}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi},\widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma},\mathbf{p})}{\partial \boldsymbol{\xi}}\right|_{\boldsymbol{\xi}=\boldsymbol{\xi}_0}=\left.\frac{\partial \mathbf{h}^{-1}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi},\mathbf{X}_s)}{\partial \boldsymbol{\xi}}\right|_{\boldsymbol{\xi}=\boldsymbol{\xi}_0}$$

Thus, the mirror parameter Jacobian is

$$\mathbf{J}_{I'}\mathbf{J}_{\mathbf{w}}\left[\mathbf{J}_{\mathbf{c}_{\boldsymbol{\xi}}}(\mathbf{0})+\mathbf{J}_{\mathbf{c}}\mathbf{J}_{\mathbf{s}}\mathbf{J}_{\mathbf{c}_{\boldsymbol{\xi}}^{-1}}(\mathbf{0})\right] \tag{A.5}$$

A.1.3 Intrinsic parameters Jacobian

The last Jacobian to be computed is the intrinsic parameters Jacobian. It is computed as the derivative

$$\left.\frac{\partial \mathbf{c}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi},\widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma},\mathbf{s}(\widehat{\mathbf{H}}\mathbf{H}(\mathbf{z}),\mathbf{c}^{-1}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi},\widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma},\mathbf{p})))}{\partial \boldsymbol{\gamma}}\right|_{\boldsymbol{\gamma}=\boldsymbol{\gamma}_0}$$

With $\mathbf{A}=\mathbf{s}(\widehat{\mathbf{H}}\mathbf{H}(\mathbf{z}),\mathbf{c}^{-1}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi},\widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma},\mathbf{p}))$ one has

$$\left.\frac{\partial \mathbf{c}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi},\widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma},\mathbf{X}_s))}{\partial \boldsymbol{\gamma}}\right|_{\boldsymbol{\gamma}=\boldsymbol{\gamma}_0} = \left.\frac{\partial \mathbf{c}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi},\widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma},\mathbf{A})}{\partial \boldsymbol{\gamma}}\right|_{\boldsymbol{\gamma}=\boldsymbol{\gamma}_0} + \mathbf{J}_{\mathbf{c}}\mathbf{J}_{\mathbf{s}}\left.\frac{\partial \mathbf{c}^{-1}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi},\widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma},\mathbf{p})}{\partial \boldsymbol{\gamma}}\right|_{\boldsymbol{\gamma}=\boldsymbol{\gamma}_0}$$

The first part is

$$\mathbf{J}_{\mathbf{c}_\gamma} = \left.\frac{\partial \mathbf{k}(\widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma},\mathbf{q}')}{\partial \boldsymbol{\gamma}}\right|_{\boldsymbol{\gamma}=\boldsymbol{\gamma}_0}$$

The second part is

$$\mathbf{J}_{\mathbf{c}_\gamma^{-1}} = \left.\frac{\partial \mathbf{c}^{-1}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi},\widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma},\mathbf{p})}{\partial \boldsymbol{\gamma}}\right|_{\boldsymbol{\gamma}=\boldsymbol{\gamma}_0} = \left.\frac{\partial \mathbf{h}^{-1}(\widehat{\boldsymbol{\xi}}+\boldsymbol{\xi},\mathbf{e})}{\partial \mathbf{e}}\right|_{\mathbf{e}=\mathbf{q}} \left.\frac{\partial \mathbf{k}^{-1}(\widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma},\mathbf{p})}{\partial \boldsymbol{\gamma}}\right|_{\boldsymbol{\gamma}=\boldsymbol{\gamma}_0}$$

Thus, the intrinsic parameters Jacobian is

$$\mathbf{J}_{I'}\mathbf{J}_{\mathbf{w}}\left[\mathbf{J}_{\mathbf{c}_\gamma}(\mathbf{0}) + \mathbf{J}_{\mathbf{c}}\mathbf{J}_{\mathbf{s}}\mathbf{J}_{\mathbf{c}_\gamma^{-1}}(\mathbf{0})\right] \tag{A.6}$$

A.2 Reference Jacobian

Like the current Jacobian, the reference Jacobian is obtained by computing the derivative of the cost function in (3.4) when $\mathbf{x} = (\mathbf{z}, \boldsymbol{\xi}, \boldsymbol{\gamma}) = \widetilde{\mathbf{x}}$.

$$\mathbf{J}(\tilde{\mathbf{x}}) = \left[\nabla_{\mathbf{x}} I'\left(\mathbf{w}\left(\widehat{\mathbf{H}}\mathbf{H}(\mathbf{z}), \widehat{\boldsymbol{\xi}}+\boldsymbol{\xi}, \widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma}, \mathbf{p}\right)\right) - I(\mathbf{p})\right]_{\mathbf{x}=\tilde{\mathbf{x}}}$$

Let us expand the function I' as follows

$$I'\left(\mathbf{w}\left(\widehat{\mathbf{H}}\mathbf{H}(\mathbf{z}), \widehat{\boldsymbol{\xi}}+\boldsymbol{\xi}, \widehat{\boldsymbol{\gamma}}+\boldsymbol{\gamma}, \mathbf{p}\right)\right) = I'(\mathbf{w}(\overline{\mathbf{H}}, \overline{\boldsymbol{\xi}}, \overline{\boldsymbol{\gamma}}, \mathbf{q}))$$

where $\mathbf{q} = \mathbf{w}^{-1}(\overline{\mathbf{H}}, \overline{\boldsymbol{\xi}}, \overline{\boldsymbol{\gamma}}, \mathbf{w}(\overline{\mathbf{H}}\widetilde{\mathbf{H}}^{-1}\mathbf{H}(\mathbf{z}), \overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{p}))$

In the optimal increment $\widetilde{\mathbf{x}}$ the exact solution is supposed to be obtained; therefore, we can verify that $\widetilde{\mathbf{H}} = \widehat{\mathbf{H}}^{-1}\overline{\mathbf{H}}$, $\widetilde{\boldsymbol{\xi}} = \overline{\boldsymbol{\xi}} - \widehat{\boldsymbol{\xi}}$ and $\widetilde{\boldsymbol{\gamma}} = \overline{\boldsymbol{\gamma}} - \widehat{\boldsymbol{\gamma}}$.

As for the current Jacobian, the derivative can be decomposed into two parts

$$\left.\frac{\partial I'(\mathbf{w}(\overline{\mathbf{H}}, \overline{\boldsymbol{\xi}}, \overline{\boldsymbol{\gamma}}, \mathbf{q}))}{\partial \mathbf{x}}\right|_{\mathbf{x}=\widetilde{\mathbf{x}}} = \left.\frac{\partial I'(\mathbf{w}(\overline{\mathbf{H}}, \overline{\boldsymbol{\xi}}, \overline{\boldsymbol{\gamma}}, \mathbf{q}))}{\partial \mathbf{q}}\right|_{\mathbf{q}=\mathbf{p}} \left.\frac{\partial \mathbf{w}^{-1}(\overline{\mathbf{H}}, \overline{\boldsymbol{\xi}}, \overline{\boldsymbol{\gamma}}, \mathbf{r}))}{\partial \mathbf{x}}\right|_{\mathbf{x}=\widetilde{\mathbf{x}}}$$

Again, the first part of the Jacobian is a (1×2) vector which can be computed

directly from the reference image data. The vector contains the gradient of the warped image. It is very important to remark that the gradient of the reference image can be computed without explicity knowing the true solution $\overline{\mathbf{x}} = (\overline{\mathbf{z}}, \overline{\boldsymbol{\xi}}, \overline{\boldsymbol{\gamma}})$. Let $\mathbf{r} = \mathbf{w}(\overline{\mathbf{H}}\widetilde{\mathbf{H}}^{-1}\mathbf{H}(\mathbf{z}), \overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{p})$. Thus, the second part of the derivative is again decomposed into two parts

$$\left.\frac{\partial \mathbf{w}^{-1}(\overline{\mathbf{H}}, \overline{\boldsymbol{\xi}}, \overline{\boldsymbol{\gamma}}, \mathbf{r})}{\partial \mathbf{x}}\right|_{\mathbf{x}=\widetilde{\mathbf{x}}} = \left.\frac{\partial \mathbf{w}^{-1}(\overline{\mathbf{H}}, \overline{\boldsymbol{\xi}}, \overline{\boldsymbol{\gamma}}, \mathbf{r})}{\partial \mathbf{q}}\right|_{\mathbf{r}=\mathbf{p}'} \left.\frac{\partial \mathbf{w}(\overline{\mathbf{H}}\widetilde{\mathbf{H}}^{-1}\mathbf{H}(\mathbf{z}), \overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{p})}{\partial \mathbf{x}}\right|_{\mathbf{x}=\widetilde{\mathbf{x}}} \tag{A.7}$$

The first part of Equation (A.7) is

$$\left.\frac{\partial \mathbf{w}^{-1}(\overline{\mathbf{H}}, \overline{\boldsymbol{\xi}}, \overline{\boldsymbol{\gamma}}, \mathbf{r})}{\partial \mathbf{r}}\right|_{\mathbf{r}=\mathbf{p}'} = \left(\left.\frac{\partial \mathbf{w}(\overline{\mathbf{H}}, \overline{\boldsymbol{\xi}}, \overline{\boldsymbol{\gamma}}, \mathbf{r})}{\partial \mathbf{r}}\right|_{\mathbf{r}=\mathbf{p}}\right)^{-1} \tag{A.8}$$

This derivative is a (2×2) matrix which represents the coordinates variation from a point $\mathbf{p}'$ in the current image I' with respect to the coordinates of a point $\mathbf{p}$ in the reference image I. This derivative will be noted $\mathbf{J}_{\widetilde{\mathbf{w}}}$. In order to compute this derivative, we will use the warping expression from (3.2), thus the derivative to be computed is

$$\left.\frac{\partial \mathbf{w}(\overline{\mathbf{H}}, \overline{\boldsymbol{\xi}}, \overline{\boldsymbol{\gamma}}, \mathbf{c}(\overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{s}(\overline{\mathbf{H}}\widetilde{\mathbf{H}}^{-1}\mathbf{H}(\mathbf{z}), \mathbf{c}^{-1}(\overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{p})))))}{\partial \mathbf{r}}\right|_{\mathbf{r}=\mathbf{p}}$$

For the purpose of saving space and to show the entire reference Jacobian in a single row, let $\widetilde{\mathbf{x}}_1 = (\overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma})$. Thus, using the chain rule for derivatives one obtains

$$\left.\frac{\partial \mathbf{w}(\overline{\mathbf{H}}, \overline{\boldsymbol{\xi}}, \overline{\boldsymbol{\gamma}}, \mathbf{r})}{\partial \mathbf{r}}\right|_{\mathbf{r}=\mathbf{p}} = \left.\frac{\partial \mathbf{c}(\widetilde{\mathbf{x}}_1, \mathbf{A})}{\partial \mathbf{A}}\right|_{\mathbf{A}=\mathbf{X}'_s} \left.\frac{\partial \mathbf{s}(\overline{\mathbf{H}}\widetilde{\mathbf{H}}^{-1}\mathbf{H}(\mathbf{z}), \mathbf{B})}{\partial \mathbf{B}}\right|_{\mathbf{B}=\mathbf{X}_s} \left.\frac{\partial \mathbf{c}^{-1}(\widetilde{\mathbf{x}}_1, \mathbf{r})}{\partial \mathbf{r}}\right|_{\mathbf{r}=\mathbf{p}}$$

We will denote these derivatives as $\mathbf{J}_{\widetilde{\mathbf{c}}}$, $\mathbf{J}_{\widetilde{\mathbf{s}}}$ and $\mathbf{J}_{\widetilde{\mathbf{c}}^{-1}}$. The first part is

$$\mathbf{J}_{\widetilde{\mathbf{c}}} = \left.\frac{\partial \mathbf{c}(\widetilde{\mathbf{x}}_1, \mathbf{A})}{\partial \mathbf{A}}\right|_{\mathbf{A}=\mathbf{X}'_s} = \left.\frac{\partial \mathbf{k}(\overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{h}(\overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \mathbf{A}))}{\partial \mathbf{A}}\right|_{\mathbf{A}=\mathbf{X}'_s}$$

Then, one obtains

$$\mathbf{J}_{\widetilde{\mathbf{c}}} = \left.\frac{\partial \mathbf{c}(\widetilde{\mathbf{x}}_1, \mathbf{A})}{\partial \mathbf{A}}\right|_{\mathbf{A}=\mathbf{X}'_s} = \left.\frac{\partial \mathbf{k}(\overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{a})}{\partial \mathbf{a}}\right|_{\mathbf{a}=\mathbf{q}'} \left.\frac{\partial \mathbf{h}(\overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \mathbf{A})}{\partial \mathbf{A}}\right|_{\mathbf{A}=\mathbf{X}'_s}$$

The second part is

$$\mathbf{J}_{\widetilde{\mathbf{s}}} = \left.\frac{\partial \mathbf{s}(\overline{\mathbf{H}}\widetilde{\mathbf{H}}^{-1}\mathbf{H}(\mathbf{z}), \mathbf{B})}{\partial \mathbf{B}}\right|_{\mathbf{B}=\mathbf{X}_s}$$

The third part is

$$\mathbf{J}_{\widetilde{\mathbf{c}}^{-1}} = \left.\frac{\partial \mathbf{c}^{-1}(\widetilde{\mathbf{x}}_1, \mathbf{r})}{\partial \mathbf{r}}\right|_{\mathbf{r}=\mathbf{p}} = \left.\frac{\partial \mathbf{h}^{-1}(\overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \mathbf{k}^{-1}(\overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{r}))}{\partial \mathbf{r}}\right|_{\mathbf{r}=\mathbf{p}}$$

Then, one obtains

$$\mathbf{J}_{\widetilde{\mathbf{c}}^{-1}} = \left.\frac{\partial \mathbf{c}^{-1}(\widetilde{\mathbf{x}}_1, \mathbf{r})}{\partial \mathbf{r}}\right|_{\mathbf{r}=\mathbf{p}} = \left.\frac{\partial \mathbf{h}^{-1}(\overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \mathbf{b})}{\partial \mathbf{b}}\right|_{\mathbf{b}=\mathbf{q}} \left.\frac{\partial \mathbf{k}^{-1}(\overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{r})}{\partial \mathbf{r}}\right|_{\mathbf{r}=\mathbf{p}}$$

This last derivative allows us to write Equation (A.8) as the inverse of the product of three Jacobians.

$$\mathbf{J}_{\widetilde{\mathbf{w}}} = \left(\mathbf{J}_{\widetilde{\mathbf{c}}}\mathbf{J}_{\widetilde{\mathbf{s}}}\mathbf{J}_{\widetilde{\mathbf{c}}^{-1}}\right)^{-1} \tag{A.9}$$

Now, we procede to computing the second part of the Equation (A.7). This derivative will be computed considering the different sets of parameters in the warping function (mirror parameter, intrinsic parameters and homography parameters). Let us start with the homography parameters.

A.2.1 Homography parameters Jacobian

The derivative to be computed is

$$\left.\frac{\partial \mathbf{c}(\overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{s}(\overline{\mathbf{H}}\widetilde{\mathbf{H}}^{-1}\mathbf{H}(\mathbf{z}), \mathbf{c}^{-1}(\overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{p})))}{\partial \mathbf{z}}\right|_{\mathbf{z}=\widetilde{\mathbf{z}}}$$

Let $\mathbf{A} = \mathbf{s}(\overline{\mathbf{H}}\widetilde{\mathbf{H}}^{-1}\mathbf{H}(\mathbf{z}), \mathbf{c}^{-1}(\overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{p}))$ and $\widetilde{\mathbf{x}}_1 = (\overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma})$. Thus, one has

$$\left.\frac{\partial \mathbf{c}(\widetilde{\mathbf{x}}_1, \mathbf{A}))}{\partial \mathbf{z}}\right|_{\mathbf{z}=\widetilde{\mathbf{z}}} = \left.\frac{\partial \mathbf{c}(\widetilde{\mathbf{x}}_1, \mathbf{A})}{\partial \mathbf{A}}\right|_{\mathbf{A}=\mathbf{X}'_s} \left.\frac{\partial \mathbf{s}(\overline{\mathbf{H}}\widetilde{\mathbf{H}}^{-1}\mathbf{H}(\mathbf{z}), \mathbf{c}^{-1}(\widetilde{\mathbf{x}}_1, \mathbf{p}))}{\partial \mathbf{z}}\right|_{\mathbf{z}=\widetilde{\mathbf{z}}}$$

The first part is the Jacobian $\mathbf{J}_{\widetilde{\mathbf{c}}}$ which has already been computed. The second part is

$$\left.\frac{\partial \mathbf{s}(\overline{\mathbf{H}}\widetilde{\mathbf{H}}^{-1}\mathbf{H}(\mathbf{z}), \mathbf{c}^{-1}(\widetilde{\mathbf{x}}_1, \mathbf{p}))}{\partial \mathbf{z}}\right|_{\mathbf{z}=\widetilde{\mathbf{z}}} = \left.\frac{\partial \mathbf{s}(\overline{\mathbf{H}}\widetilde{\mathbf{H}}^{-1}\mathbf{H}(\mathbf{z}), \mathbf{c}^{-1}(\widetilde{\mathbf{x}}_1, \mathbf{p}))}{\partial \mathbf{H}(\mathbf{z})}\right|_{\mathbf{H}(\mathbf{z})=\widetilde{\mathbf{H}}} \left.\frac{\partial \overline{\mathbf{H}}\widetilde{\mathbf{H}}^{-1}\mathbf{H}(\mathbf{z})}{\partial \mathbf{z}}\right|_{\mathbf{z}=\widetilde{\mathbf{z}}}$$

These derivatives end up with the homography parameters Jacobian. We will note them as $\mathbf{J}_{\widetilde{\mathbf{s}}_\mathbf{H}}$ and $\mathbf{J}_{\widetilde{\mathbf{H}}_\mathbf{z}}$. Thus, the homography parameters Jacobian can be written as

$$\mathbf{J}_I \mathbf{J}_{\widetilde{\mathbf{w}}} \mathbf{J}_{\widetilde{\mathbf{c}}} \mathbf{J}_{\widetilde{\mathbf{s}}_\mathbf{H}} \mathbf{J}_{\widetilde{\mathbf{H}}_\mathbf{z}}(\widetilde{\mathbf{x}}) \tag{A.10}$$

A.2.2 Mirror parameter Jacobian

Now, we will compute the mirror parameter Jacobian. Similarly to the last Jacobian, the derivative to be computed is

$$\left.\frac{\partial \mathbf{c}(\overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{s}(\overline{\mathbf{H}}\widetilde{\mathbf{H}}^{-1}\mathbf{H}(\mathbf{z}), \mathbf{c}^{-1}(\overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{p})))}{\partial \boldsymbol{\xi}}\right|_{\boldsymbol{\xi}=\widetilde{\boldsymbol{\xi}}}$$

Again, let $\mathbf{A} = \mathbf{s}(\overline{\mathbf{H}}\widetilde{\mathbf{H}}^{-1}\mathbf{H}(\mathbf{z}), \mathbf{c}^{-1}(\overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{p}))$ and $\widetilde{\mathbf{x}}_1 = (\overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma})$ as the previous Jacobian. Then, one has

$$\left.\frac{\partial \mathbf{c}(\widetilde{\mathbf{x}}_1, \mathbf{X}_s))}{\partial \boldsymbol{\xi}}\right|_{\boldsymbol{\xi}=\widetilde{\boldsymbol{\xi}}} = \left.\frac{\partial \mathbf{c}(\overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{A})}{\partial \boldsymbol{\xi}}\right|_{\boldsymbol{\xi}=\widetilde{\boldsymbol{\xi}}} + \mathbf{J}_{\widetilde{\mathbf{c}}} \mathbf{J}_{\widetilde{\mathbf{s}}} \left.\frac{\partial \mathbf{c}^{-1}(\overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{p})}{\partial \boldsymbol{\xi}}\right|_{\boldsymbol{\xi}=\widetilde{\boldsymbol{\xi}}}$$

The first part is

$$\mathbf{J}_{\widetilde{\mathbf{c}}_{\boldsymbol{\xi}}} = \left.\frac{\partial \mathbf{c}(\overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{A})}{\partial \boldsymbol{\xi}}\right|_{\boldsymbol{\xi}=\widetilde{\boldsymbol{\xi}}} = \left.\frac{\partial \mathbf{k}(\overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{a})}{\partial \mathbf{a}}\right|_{\mathbf{a}=\mathbf{q}'} \left.\frac{\partial \mathbf{h}(\overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \mathbf{X}_s)}{\partial \boldsymbol{\xi}}\right|_{\boldsymbol{\xi}=\widetilde{\boldsymbol{\xi}}}$$

The last part is

$$\mathbf{J}_{\widetilde{\mathbf{c}}_{\boldsymbol{\xi}}^{-1}} = \left.\frac{\partial \mathbf{c}^{-1}(\overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{p})}{\partial \boldsymbol{\xi}}\right|_{\boldsymbol{\xi}=\widetilde{\boldsymbol{\xi}}} = \left.\frac{\partial \mathbf{h}^{-1}(\widehat{\boldsymbol{\xi}} + \boldsymbol{\xi}, \mathbf{X}_s)}{\partial \boldsymbol{\xi}}\right|_{\boldsymbol{\xi}=\widetilde{\boldsymbol{\xi}}}$$

Thus, the mirror parameter Jacobian is

$$\mathbf{J}_I \mathbf{J}_{\widetilde{\mathbf{w}}} \left[\mathbf{J}_{\widetilde{\mathbf{c}}_{\boldsymbol{\xi}}}(\widetilde{\mathbf{x}}) + \mathbf{J}_{\widetilde{\mathbf{c}}} \mathbf{J}_{\widetilde{\mathbf{s}}} \mathbf{J}_{\widetilde{\mathbf{c}}_{\boldsymbol{\xi}}^{-1}}(\widetilde{\mathbf{x}})\right] \tag{A.11}$$

A.2.3 Intrinsic parameters Jacobian

The last Jacobian to be computed is the intrinsic parameters Jacobian. It is computed with the next derivative

$$\left.\frac{\partial \mathbf{c}(\overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{s}(\overline{\mathbf{H}}\widetilde{\mathbf{H}}^{-1}\mathbf{H}(\mathbf{z}), \mathbf{c}^{-1}(\overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{p})))}{\partial \boldsymbol{\gamma}}\right|_{\boldsymbol{\gamma}=\widetilde{\boldsymbol{\gamma}}}$$

Let $\widetilde{\mathbf{x}}_1$ and $\mathbf{A}$ be the same than the previous section. Thus one has

$$\left.\frac{\partial \mathbf{c}(\widetilde{\mathbf{x}}_1, \mathbf{X}_s))}{\partial \boldsymbol{\gamma}}\right|_{\boldsymbol{\gamma}=\widetilde{\boldsymbol{\gamma}}} = \left.\frac{\partial \mathbf{c}(\overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{A})}{\partial \boldsymbol{\gamma}}\right|_{\boldsymbol{\gamma}=\widetilde{\boldsymbol{\gamma}}} + \mathbf{J}_{\widetilde{\mathbf{c}}} \mathbf{J}_{\widetilde{\mathbf{s}}} \left.\frac{\partial \mathbf{c}^{-1}(\overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{p})}{\partial \boldsymbol{\gamma}}\right|_{\boldsymbol{\gamma}=\widetilde{\boldsymbol{\gamma}}}$$

The first part is

$$\mathbf{J}_{\widetilde{\mathbf{c}}_{\boldsymbol{\gamma}}} = \left.\frac{\partial \mathbf{k}(\overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{q}')}{\partial \boldsymbol{\gamma}}\right|_{\boldsymbol{\gamma}=\widetilde{\boldsymbol{\gamma}}}$$

The second part is

$$\mathbf{J}_{\widetilde{\mathbf{c}}_{\boldsymbol{\gamma}}^{-1}} = \left.\frac{\partial \mathbf{c}^{-1}(\overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{p})}{\partial \boldsymbol{\gamma}}\right|_{\boldsymbol{\gamma}=\widetilde{\boldsymbol{\gamma}}} = \left.\frac{\partial \mathbf{h}^{-1}(\overline{\boldsymbol{\xi}} - \widetilde{\boldsymbol{\xi}} + \boldsymbol{\xi}, \mathbf{e})}{\partial \mathbf{e}}\right|_{\mathbf{e}=\mathbf{q}} \left.\frac{\partial \mathbf{k}^{-1}(\overline{\boldsymbol{\gamma}} - \widetilde{\boldsymbol{\gamma}} + \boldsymbol{\gamma}, \mathbf{p})}{\partial \boldsymbol{\gamma}}\right|_{\boldsymbol{\gamma}=\widetilde{\boldsymbol{\gamma}}}$$

Thus, the intrinsic parameters Jacobian is

$$\mathbf{J}_I \mathbf{J}_{\widetilde{\mathbf{w}}} \left[\mathbf{J}_{\widetilde{\mathbf{c}}_{\boldsymbol{\gamma}}}(\widetilde{\mathbf{x}}) + \mathbf{J}_{\widetilde{\mathbf{c}}} \mathbf{J}_{\widetilde{\mathbf{s}}} \mathbf{J}_{\widetilde{\mathbf{c}}_{\boldsymbol{\gamma}}^{-1}}(\widetilde{\mathbf{x}})\right] \tag{A.12}$$

Appendix B

Theoretical demonstrations

B.1 Proof of Property 1

Proof of Property 1. From relation $A_2\sqrt{A_4} = B_2\sqrt{B_4}$ one deduces $A_2^2A_4 = B_2^2B_4$. Since A_2, B_2 are irreducible, the relation $A_2^2A_4 = B_2^2B_4$ indicates that either B_2 divides A_2 or B_4 divides A_2. Let us consider these two cases.

- If B_2 divides A_2, then there exists a constant $\eta \neq 0$ such that $A_2 = \eta B_2$. From here, it is straightforward to deduce that $\sqrt{B_4} = \eta\sqrt{A_4}$.
- If B_4 divides A_2, then there exists a second order polynomial B_{2a} such that $B_4 = A_2B_{2a}$. From here, one deduces from the relation $A_2^2A_4 = B_2^2B_4$ that $A_2A_4 = B_2^2B_{2a}$. This latter relation and the facts that A_2, B_2 are irreducible imply that either B_2 divides A_2 or B_{2a} divides A_2. Let us study these two cases:

  1) If B_2 divides A_2, then analogously to what have been shown previously one easily deduces the desired property.

  2) If B_{2a} divides A_2, then there exists a constant α such that $B_{2a} = \alpha A_2$. From here, one deduces that $B_4 = A_2B_{2a} = \alpha A_2^2$ which implies that $\sqrt{B_4}(= \pm\sqrt{\alpha}A_2)$ is a polynomial. However, this is a contradiction with the assumption made in Property 1.

All cases lead to the desidere property. □

B.2 Proof of Property 2

Proof of Property 2. Let us prove Property 2 by contradiction. Assume that $P_m \neq 0$, $P_n \neq 0$. Powered by 2 relation $P_m = P_n\sqrt{P_l}$ one obtains $P_m^2 = P_n^2P_l$. From this expression one deduces that P_m^2 is divisible by P_n^2 thus P_m is divisible by P_n. As a consequence $\sqrt{P_l}$ musbe a polynomial. The recently contradiction indicates that $P_n \equiv P_m \equiv 0$. □

B.3 Proof of Property 4

Proof of Property 4. Let us consider two possible cases:

Case 1: Assume that $a_4 = 0$. Then, $A_2 = a_1x + a_2y + a_3$ which implies that

A_2 is irreducible.

Case 2: Consider the case $a_4 \neq 0$ and proceed the proof by contradiction. Assume that A_2 is an irreducible polynomial; so that$A_2 = (\alpha_1 x + \alpha_2 y + \alpha_3)(\beta_1 x + \beta_2 y + \beta_3)$ for some constants $\alpha_{1,2,3}$, $\beta_{1,2,3}$. From the coefficients in xy, x, y, x^2, y^2 and the constant term of A_2 one deduces the following relations

$$\alpha_1\beta_2 + \alpha_2\beta_1 = 0 \tag{B.1}$$

$$\alpha_1\beta_3 + \alpha_3\beta_1 = a_1 \tag{B.2}$$

$$\alpha_1\beta_3 + \alpha_3\beta_2 = a_2 \tag{B.3}$$

$$\alpha_3\beta_3 = a_3 \tag{B.4}$$

$$\alpha_1\beta_1 = a_4 \tag{B.5}$$

$$\alpha_2\beta_2 = a_4 \tag{B.6}$$

Then, from (B.1) one deduces $\beta_2 = (-\alpha_2\beta_1)/\alpha_1$. From (B.5) and (B.6) one has $\alpha_1\beta_1 = \alpha_2\beta_2$. From this one has $\alpha_1^2\beta_1 = -\alpha_2^2\beta_1$. If $\beta_1 \neq 0$ one deduces that $\alpha_1 = \alpha_2 = 0$, therefore $A_2 = \alpha_3(\beta_1 x + \beta_2 y + \beta_3)$ is irreducible. If $\beta_1 = 0$, then $a_4 = 0$ which implies that A_2 is irreducible. □

B.4 Proof of Lemma 1

Proof of Lemma 1. If $\sqrt{B_3 + B_4}$ is a poylomial one verifies $B_3 + B_4 = (a_1x^2 + a_2y^2 + a_3xy + a_4x + a_5y + a_6)^2$ with some constants $a_{1,2,3,4,5,6}$. From (4.29) one has

$$B_3 + B_4 = \sum_{i=1}^{3}((P_1^i) + (P_2^i))^2$$

From (4.27) one has $P_1^i = h_{i1}x + h_{i2}y + h_{i3}$ and $P_2^i = h_{i1}x + h_{i2}y + h_{i3}(-x^2 - y^2)$. Therefore,

$$\sum_{i=1}^{3}\left(2h_{i1}x + 2h_{i2}y + h_{i3}(1 - x^2 - y^2)\right)^2 = (a_1x^2 + a_2y^2 + a_3xy + a_4x + a_5y + a_6)^2$$

then

$$\begin{aligned}
&\sum_{i=1}^{3} 4h_{i1}^2x^2 + \sum_{i=1}^{3} 4h_{i2}^2y^2 + \sum_{i=1}^{3} h_{i3}^2(1 + x^4 + y^4 + 2x^2y^2 - 2x^2 - 2y^2) \\
&+ \sum_{i=1}^{3} 4h_{i1}h_{i2}xy + \sum_{i=1}^{3} 4h_{i1}h_{i3}(-x^3 - xy^2 + x) + \sum_{i=1}^{3} 4h_{i2}h_{i3}(-x^2y - y^3 + y) \\
&= a_1^2x^4 + a_2^2y^4 + a_3^2x^2y^2 + a_4^2x^2 + a_5^2y^2 + a_6^2 + 2a_1a_2x^2y^2 + 2a_1a_3x^3y + 2a_1a_4x^3 \\
&+2a_1a_5x^2y + 2a_1a_6x^2 + 2a_2a_3xy^3 + 2a_2a_4xy^2 + 2a_2a_5y^3 + 2a_2a_6y^2 + 2a_3a_4x^2y \\
&+2a_3a_5xy^2 + 2a_3a_6xy + 2a_4a_5xy + 2a_4a_6x + 2a_5a_6y
\end{aligned} \tag{B.7}$$

From the coefficients of x^4, y^4 and the constant term of (B.7) one has

$$a_1^2 = a_2^2 = a_6^2 = \sum_{i=1}^{3} h_{i3}^2 \tag{B.8}$$

Since $\det(\mathbf{H}) \neq 0$, $\sum_{i=1}^{3} h_{i3}^2$. Therefore, one deduces that $|a_1| = |a_2| = |a_6|$.

From the coefficients of x and y of (B.7) one has

$$a_4a_6 = 2\sum_{i=1}^{3} h_{i1}h_{i3} \tag{B.9}$$

$$a_5a_6 = 2\sum_{i=1}^{3} h_{i2}h_{i3} \tag{B.10}$$

From the coefficients of x^2 and y^2 of (B.7) one has

$$a_4^2 + 2a_1a_6 = -2\sum_{i=1}^{3} h_{i3}^2 + 4\sum_{i=1}^{3} h_{i1}^2 \tag{B.11}$$

$$a_5^2 + 2a_2a_6 = -2\sum_{i=1}^{3} h_{i3}^2 + 4\sum_{i=1}^{3} h_{i2}^2 \tag{B.12}$$

From the coefficients of x^3 and y^3 of (B.7) one has

$$a_1a_4 = -2\sum_{i=1}^{3} h_{i1}h_{i3} \tag{B.13}$$

$$a_2a_5 = -2\sum_{i=1}^{3} h_{i2}h_{i3} \tag{B.14}$$

From the coefficients of x^3y and xy^3 of (B.7) one has

$$a_1 a_3 = 0 \tag{B.15}$$

$$a_2 a_3 = 0 \tag{B.16}$$

From (B.15) and (B.16) one deduces that $a_3 = 0$, since $a_1 \neq 0$ and $a_2 \neq 0$ (proved previously).

Looking for x^2y^2 of (B.7) one has

$$a_1 a_2 = \sum_{i=1}^{3} h_{i3}^2 \tag{B.17}$$

From (B.8) and (B.17) one deduces $a_1 = a_2 \neq 0$.

From the coefficients of xy of (B.7) one has

$$a_4 a_5 = 2 \sum_{i=1}^{3} h_{i1} h_{i2} \tag{B.18}$$

From the coefficients of xy^2 and x^2y of (B.7) one has

$$a_2 a_4 = -2 \sum_{i=1}^{3} h_{i1} h_{i3} \tag{B.19}$$

$$a_1 a_5 = -2 \sum_{i=1}^{3} h_{i2} h_{i3} \tag{B.20}$$

Now, let us prove that both a_4,a_5 are equal to zero or different from zero. Let us consider two cases:

1) **Case** $a_4 \neq 0$: From (B.9), (B.19) one has $a_4 a_6 = -a_4 a_2$. Since $a_4 \neq 0$, one has $a_6 = -a_2 \neq 0$. Thus, from (B.8) one deduces that

$$a_2 a_6 = -\sum_{i=1}^{3} h_{i3}^2 \tag{B.21}$$

From (B.21) and (B.12) one obtains

$$a_5^2 = 4 \sum_{i=1}^{3} h_{i2}^2 \neq 0 \tag{B.22}$$

This implies that $a_5 \neq 0$.

2) **Case** $a_5 \neq 0$: From (B.10), (B.20) one has $a_5 a_6 = -a_5 a_1$. Since $a_5 \neq 0$, one has $a_6 = -a_1 \neq 0$. Thus, from (B.8) one deduces that

$$a_1 a_6 = -\sum_{i=1}^{3} h_{i3}^2 \tag{B.23}$$

From (B.23) and (B.11) one obtains

$$a_4^2 = 4\sum_{i=1}^{3} h_{i1}^2 \neq 0 \tag{B.24}$$

This implies that $a_4 \neq 0$.

These two cases conclude the mentioned above remark that both a_4, a_5 are equal to zero or both different from zero.

Let us assume that $a_4 \neq 0$ and $a_5 \neq 0$ and show that this case is impossible. Relations (B.21), (B.22), (B.23), (B.24) are valid. Besides, one has $a_1 = a_2 = -a_6$.

From (B.8), (B.24), and (B.13) one has

$$\frac{1}{4}a_1^2 a_4^2 = \sum_{i=1}^{3} h_{i1}^2 h_{i3}^2 = \left(\sum_{i=1}^{3} h_{i1}^2 h_{i3}^2\right)^2 \tag{B.25}$$

Besides, one verifies that for all $x_1, x_2, x_3, y_1, y_2, y_3 \in \mathbb{R}$,

$$(x_1^2 + x_2^2 + x_3^2)(y_1^2 + y_2^2 + y_3^2) \geq (x_1 y_1 + x_2 y_2 + x_3 y_3)^2$$

Where equality occurs only if there exists a constant k such that $[x_1,\ x_2,\ x_3]^\top = k[y_1,\ y_2,\ y_3]^\top$. This property and (B.25) indicates that there exists a constant k such that $[h_{11},\ h_{21},\ h_{31}]^\top = k[h_{13},\ h_{23},\ h_{33}]^\top$. But this implies that $\det(\mathbf{H}) = 0$. This resulting contradiction with the fact that $\det(\mathbf{H}) = 1$ ensures that $a_4 = a_5 = 0$.

Now, we consider the case that $a_4 = a_5 = 0$. From (B.9), (B.18) and (B.11) one has

$$\sum_{i=1}^{3} h_{i1} h_{i3} = 0 \tag{B.26}$$

$$\sum_{i=1}^{3} h_{i1} h_{i2} = 0 \tag{B.27}$$

$$2a_1 a_6 = -2\sum_{i=1}^{3} h_{i3}^2 + 4\sum_{i=1}^{3} h_{i1}^2 \tag{B.28}$$

Since $\sum_{i=1}^{3} h_{i1}$, from (B.28) and (B.8) one easily deduces

$$a_1 = a_6 \tag{B.29}$$

Then, from (B.28) and (B.29) one has

$$\sum_{i=1}^{3} h_{i3}^2 = \sum_{i=1}^{3} h_{i1}^2 \tag{B.30}$$

In summary, from (B.26), (B.10), (B.27), (B.8), (B.30) and (B.12) one has

$$\begin{cases} \sum_{i=1}^{3} h_{i1}h_{i3} = 0 \\ \sum_{i=1}^{3} h_{i2}h_{i3} = 0 \\ \sum_{i=1}^{3} h_{i1}h_{i2} = 0 \\ a_1 = a_2 = a_6 \\ \sum_{i=1}^{3} h_{i1}^2 = \sum_{i=1}^{3} h_{i2}^2 = \sum_{i=1}^{3} h_{i3}^2 \end{cases}$$

Let us denote $\mathbf{u} = (h_{11}; h_{21}; h_{31})$, $\mathbf{v} = (h_{12}; h_{22}; h_{32})$, and $\mathbf{w} = (h_{13}; h_{23}; h_{33})$. One deduces

$$\begin{cases} \|\mathbf{u}\| = \|\mathbf{v}\| = \|\mathbf{w}\| \\ < \mathbf{u}, \mathbf{v} >= 0 \\ < \mathbf{v}, \mathbf{w} >= 0 \\ < \mathbf{u}, \mathbf{w} >= 0 \end{cases} \tag{B.31}$$

Which means that $\mathbf{u} \perp \mathbf{v}$, $\mathbf{u} \perp \mathbf{w}$, and $\mathbf{v} \perp \mathbf{v}$ and $\det(\mathbf{H}) = (\mathbf{u} \times \mathbf{v})^\top \mathbf{w} = \|\mathbf{u}\|^3 = 1$. So, one deduces that $\|\mathbf{u}\| = \|\mathbf{v}\| = \|\mathbf{w}\| = 1$ which in this implies that $\mathbf{H} \in \mathbb{SO}(3)$. □

Bibliography

[Ali01] D.G. Aliaga. Accurate catadioptric calibration for real-time pose estimation in room-size environments. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 1, pages 127–134 vol.1, 2001. 9, 24

[BA05] J.P. Barreto and H. Araujo. Geometric properties of central catadioptric line images and their application in calibration. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(8):1327–1333, Aug. 2005. 9, 23, 37, 98

[Bar03] João P. Barreto. *General Central Projection Systems, modeling, calibration and visual servoing*. PhD thesis, Department of electrical and computer engineering, 2003. 16, 19

[BK01] R. Benosman and S. B. Kang. A brief historical perspective on panorama. In *Panoramic Vision*, pages 5–20. Apr 2001. 7

[BMH02] João Barreto, Frederick Martin, and Radu P. Horaud. Visual servoing/tracking using central catadioptric cameras. In *International Symposium on Experimental Robotics*, Advanced Robotics Series, July 2002. 32

[BP02] Hynek Bakstein and Tomás Padjla. Panoramic mosaicing with a 180 degrees field of view lens. In *In Proceedings of the IEEE Workshop on Omndirectional Vision*, pages 60–67, 2002. 9, 64

[BPS+08] Yalin Bastanlar, Luis Puig, Peter Sturm, Josechu Guerrero, and João Barreto. Dlt-like calibration of central catadioptric cameras. In *Proceedings of the Workshop on Omnidirectional Vision, Camera Networks and Non-Classical Cameras, Marseille, France*, Oct 2008. 9, 20, 21

[BS] A. Basu and D. Southwell. Omni-directional sensor for pipe inspection. 4:3107–3112. 7

[BV08] C. Demonceaux Bazin and P. Vasseur. Automatic calibration of catadioptric cameras in urban environment. In *IROS*, 2008. 9, 64, 66

[CBDD00] C. Cauchois, E. Brassart, L. Delahoche, and T. Delhommelle. Reconstruction with the calibrated syclop sensor. In *IEEE International Conference on Intelligent Robots and Systems*, pages 1493–1498, Kagawa University, Takamatsu, Japan, October- November 2000. 9, 20, 22

[Cha76] Jeffrey R. Charles. Converting panoramas to circular images and vice versa without a computer. In *IEEE Transactions on Robotics and Automation*, volume http://www.eclipsechaser.com/eclink/astrotec/panconv.htm, 1976. 11

[CMEM07] J. Courbon, Y. Mezouar, L. Eck, and P. Martinet. A generic fisheye camera model for robotic applications. In *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*, pages 1683–1688, 29 2007-Nov. 2 2007. 37

[CS97] J.S. Chahl and M.V. Srinivasan. Reflective surfaces for panoramic imaging. In *Applied Optics*, volume 36, pages 8275–8285, 1997. 11

[DGSW00] A. K. Jain D. Gutchess and Sei-Wang. Automatic surveillance using omnidirectional and active cameras. In *In Asian Conference on Computer Vision (ACCV)*, January 2000. 7

[DS37] R. Descartes and D. Smith. *The geometry of René Descartes*. Originally published in Discours de la Methode. Dover Publ.: New York, 1637. 11

[DSR96] M. Fiala D. Southwell, A. Basu and J. Reyda. Panoramic stereo. In *In International Conference on Pattern Recognition*, pages 378–382, 1996. 12

[DWW07] Xiao-Ming Deng, Fu-Chao Wu, and Yi-Hong Wu. An easy calibration method for central catadioptric cameras. *Acta Automatica Sinica*, 33(8):801 – 808, 2007. 9, 20, 98

[EMW01] E. Pagello E. Menegatti and M. Wright. A new omnidirectional vision sensor for the spatial semantic hierarchy. *In IEEE/ASME AIM*, pages 93–98, 2001. 7

[Fau93] O. Faugeras. *Three-dimensional computer vision - a geometric viewpoint*. MIT Press, 1993. 13, 14, 69

[FB81] Martin A. Fischler and Robert C. Bolles. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM*, 24(6):381–395, 1981. 30

[FTB02] J. Fabrizio, J.-P. Tarel, and R. Benosman. Calibration of panoramic catadioptric sensors made easier. In *Omnidirectional Vision, 2002. Proceedings. Third Workshop on*, pages 45–52, 2002. 9, 64, 66

[GD00] Christopher Geyer and Konstantinos Daniilidis. A unifying theory for central panoramic systems and practical applications. In *European Conference on Computer Vision*, pages 445–461, 2000. 16

[GD01] C. Geyer and K. Daniilidis. Structure and motion from uncalibrated catadioptric views. *In Proceedings of CVPR'01*, pages 279–286, 2001. 9, 64

[GD02] C. Geyer and K. Daniilidis. Paracatadioptric camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):687–695, 2002. 9, 22, 23

[Geb98] Tomas Geb. Real-time panospheric image dewarping and presentation for remote mobile robot control. In *http://grok.ecn.uiowa.edu/oldIndex.html*, 1998. 11

[Gey03] Christopher Geyer. *Catadioptric Projective Geometry : theory and applications*. PhD thesis, University of Pennsylvania, 2003. 16

[GN98] J. Gluckman and K. Nayar. Ego-motion and omnidirectional cameras. In *In Proceedings of the IEEE ICCV'98*, pages 999–1005, 1998. 9, 64

[Gre85] P. Greguss. The tube peeper: A new concept in endoscopy. *In Optics and Laser Technologies*, pages 41–45, 1985. 7

[GSB09] Simone Gasparini, Peter Sturm, and João Barreto. Plane-based calibration of central catadioptric cameras. In *Proceedings of the International Conference on Computer Vision, Kyoto, Japan*, 2009. to appear. 9, 20, 21

[Ham97] Francis Hamit. New video and still cameras provide a global roaming viewpoint. In *Advance Imaging*, pages 50–52, 1997. 12

[HAMAM05] H. Hadj-Abdelkader, Y. Mezouar, N. Andreff, and P. Martinet. 2 1/2 d visual servoing with central catadioptric cameras. In *IEEE International Conference on Intelligent Robots and Systems*, 2005. 32, 38

[HB99] A.R. Hicks and R. Bajcsy. Reflective surfaces as computational sensors. In *In the second IEEE Workshop on Perception for Mobile Agents (CVPR'99)*, 1999. 12

[HZ00] Richard Hartley and Andrew Zisserman. *Multiple View geometry in Computer vision*. Cambridge university press, 2000. 13, 14, 17, 30

[JO02] Valdir Grassi Jun Okamoto. Visual servo control of a mobile robot using omnidirectional vision. In *Proc. of Mechatronics*, University of Twente, pages 24–26, June 2002. 32, 33

[Kan00] Sing Bing Kang. Catadioptric self-calibration. In *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, volume 1, pages 201–207 vol.1, 2000. 9, 64

[LK81] B. D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. *International Joint Conference on Artificial Intelligence*, pages 674–679, 1981. 31, 32

[LLF05] V. Lepetit, P. Lagger, and P. Fua. Randomized trees for real-time keypoint recognition. In *Computer Vision and Pattern Recognition*, June 2005. 30

[Low04] David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 2(60):91–110, 2004. 30

[LSEM08] J.F. Layerle, X. Savatier, J.Y. Ertaud, and E.M. Mouaddib. An innovative catadioptric vision device for the driver assistant systems. In *International Conference on Vehicle and Infrastructure Safety Improvement in Adverse Conditions and Night Driving (VISION'08)*, October 2008. 12

[Mal04] E. Malis. Improving vision-based control using efficient second-order minimization techniques. In *IEEE International Conference on Robotics and Automation*, 2004. 40

[MBMR06] C. Mei, S. Benhimane, E. Malis, and P. Rives. Homography-based tracking for central catadioptric cameras. In *IEEE International Conference on Intelligent Robots and Systems*, October 2006. 9, 20, 21, 32, 34, 38, 39, 71, 76, 81, 82, 98

[MBMR08] C. Mei, S. Benhimane, E. Malis, and P. Rives. Efficient homography-based tracking and 3-d reconstruction for single-viewpoint sensors. *IEEE Transactions on Robotics*, 24(6):1352–1364, Dec. 2008. 33, 34, 40, 41, 42, 45, 46, 47, 50, 51, 52, 53, 55, 56

[MC02] J. Matas and O. Chum. Randomized ransac with t(d,d) test. In *British Machine Vision Conference*, 2002. 30

[MP04] Branislav Micusik and Tomas Pajdla. Para-catadioptric camera auto-calibration from epipolar geometry, 2004. 9, 64

[MR07] C. Mei and P. Rives. Single view point omnidirectional camera calibration from planar grids. In *IEEE International Conference on Robotics and Automation*, April 2007. 37

[MSF07] O. Morel, R. Seulin, and D. Fofi. Catadioptric camera calibration by polarization imaging. In *IbPRIA07*, pages II: 396–403, 2007. 9, 26

[MV07] E. Malis and M. Vargas. Deeper understanding of the homography decomposition for vision-based control. In *Research report*, pages 37–38. September 2007. 93

[Nay97] Shree K. Nayar. Catadioptric omndirectional camera. In *International Conference On Computer Vision and Pattern Recognition*, pages 482–488, 1997. 12

[Num02] *Numerical Recipes in C++: The Art of Scientific Computing*. Cambridge University Press, February 2002. 64

[PLF05] J. Pilet, V. Lepetit, and P. Fua. Real-time non-rigid surface detection. In *Computer Vision and Pattern Recognition*, June 2005. 30

[RCS02] A. Gupta J. C. I. Tashev L. wei He A. Colburn Z. Zhang Z. Liu R. Cutler, Y. Rui and S. Silverberg. Distributed meetings: A meeting capture and broadcasting system. In *In ACM Multimedia*, 2002. 7

[Ree70] D. W. Rees. Panoramic television viewing system. United States Patent No. 3, 505, 465, April 1970. 11

[RPFS63] R. B. Leighton R. P. Feynman and M. Sands. *The Feynman Lectures on Physics*, volume 1. Addison-Wesley, 1963. 11

[SB08] Peter Sturm and João Barreto. General imaging geometry for central catadioptric cameras. In *Proceedings of the 10th European Conference on Computer Vision, Marseille, France*, volume 4, pages 609–622. Springer, oct 2008. 9, 20, 21

[SFZ00] G. Simon, A. Fitzgibbon, and A. Zisserman. Markerless tracking using planar structures in the scene. In *IEEE International Symposium on Mixed and Augmented Reality*, 2000. 30

[SMS06] D. Scaramuzza, A. Martinelli, and R. Siegwart. A toolbox for easy calibrating omnidirectional cameras. In Proceedings of IEEE International Conference on Intelligent Robots and Systems, 2006. 9, 20, 21

[SR04] Peter Sturm and Srikumar Ramalingam. A generic concept for camera calibration. In *Proceedings of the European Conference on Computer Vision, Prague, Czech Republic*, volume 2, pages 1–13. Springer, May 2004. 9

[SRL06] Peter Sturm, Srikumar Ramalingam, and Suresh Lodha. On calibration, structure from motion and multi-view geometry for generic camera models. In Kostas Daniilidis and Reinhard Klette, editors, *Imaging Beyond the Pinhole Camera*, volume 33 of *Computational Imaging and Vision Series*. Springer, 2006. 26

[ST94a] Jianbo Shi and C. Tomasi. Good features to track. In *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994 IEEE Computer Society Conference on*, pages 593–600, June 1994. 64

[ST94b] Jianbo Shi and Carlo Tomasi. Good features to track. In *IEEE Conference on Computer Vision and Pattern Recognition*, 1994. 31, 32

[TG04] Mohan Trivedi Tarak Gandhi. Motion analysis for event detection and tracking with a mobile omnidirectional camera. 10(2):131–143, August 2004. 32

[TKY98] H. Iwasa H. Takemura T. Kawanishi, K. Yamazawa and N. Yokoya. Generation of high-resolution stereo panoramic images by omnidirectional imaging sensor using hexagonal pyramidal mirros. In *14th International Conference on Pattern Recognition*, volume 1, pages 485–489, August 1998. 11

[TMdCM02] B. Tordoff, W. Mayol, T. de Campos, and D. Murray. Head pose estimation for wearable robot control. In *British Machine Vision Conference*, 2002. 30

[Tsa89] Roger Tsai. Synopsis of recent progress on camera calibration for 3d machine vision. pages 147–159, 1989. 69

[Vas04] Mouaddib Vasseur. Central catadioptric line detection. In *In: Proceedings of the British Machine Vision Conference*. BMVA Press, 2004. 9, 24, 98

[WH05] Yihong Wu and Zhanyi Hu. Geometric invariants and applications under catadioptric camera model. In *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, volume 2, pages 1547–1554 Vol. 2, Oct. 2005. 9, 24

[WLH06] Yihong Wu, Youfu Li, and Zhanyi Hu. Easy calibration for para-catadioptric-like camera. In *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*, pages 5719–5724, Oct. 2006. 9, 23, 98

[XY05] Hongbin Zha Xianghua Ying. Linear catadioptric camera calibration from spheres. In *OMNIVIS*, 2005. 9, 25

[Yag99] Yasushi Yagi. Omnidirectional sensing and its applications. *IEICE Trans, on Information and Systems*, E82-D(3):568–579, 1999. 7

[YH03] Xianghua Ying and Zhanyi Hu. Catadioptric camera calibration using geometric invariants. *Computer Vision, IEEE International Conference on*, 2:1351, 2003. 9, 25, 98

[YY00] Y. Yamazawa, K.and Yagi and M. Yachida. 3d line segment reconstruction by using hyperomni vision and omnidirectional hough transforming. In *ICPR*, 2000. 12

[YYY95] Y. Nishizawa Y. Yagui and M. Yachida. Map-based navigation for a mobile robot with omnidirectional images sensor copis. In *IEEE Transaction on Robotics and AUtomation*, 1995. 12

[YZ05] Xianghua Ying and Hongbin Zha. Simultaneously calibrating catadioptric camera and detecting line features using hough transform. In *Intelligent Robots and Systems, 2005. (IROS 2005). 2005 IEEE/RSJ International Conference on*, pages 412–417, Aug. 2005. 9, 23, 66, 98

Direct Self-Calibration of Central Catadioptric Omnidirectional Cameras

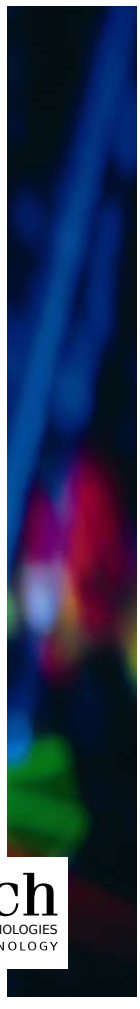
Abstract:

Calibrating a camera means determining the geometric properties of the imaging process, i.e., the transformation that maps a three-dimensional point, expressed with respect to a reference frame, onto its two-dimensional image whose coordinates are expressed in pixel units. This process is required when recovering 3D information. More precisely, we have to know the translation and rotation of the visual sensor with respect to the rest of the frame system (extrinsic parameters), and the different parameters of the lens, such as focal length, magnitude factors, optical center retinal location (intrinsic parameters).

Although the camera calibration problem is well understood, no method allowing the robust direct on-line self-calibration for any central omnidirectional camera is known. Existing self-calibration techniques attempt to calibrate from point correspondences, lines, circles or a specific camera motion. Even though interesting results can be achieved, self-calibration still suffers from some limitations such as small number of feature points, difficult detection of lines, undesirable camera motion and taking into account a specific mirror. Consequently, the aim of this thesis is to propose a new algorithm that overcomes these limitations and can be adopted by any robotic application or by any other practical implementation in which the calibration process is not straightforward; this algorithm works directly with the image intensity, makes the minimum of assumptions about the particular structure of the scene being viewed, stays valid for all central catadioptric systems and needs no prior knowledge about extrinsic and intrinsic parameters.

Also, part of this thesis is dedicated to formalize the uniqueness of the solution for the calibration problem of central catadioptric omnidirectional cameras. For the greatest part of the work on omnidirectional camera calibration it has been observed that, in the case of a non-planar mirror, two images acquired from different points of view suffice to calibrate an omnidirectional camera. However, to our knowledge, no theoretical proof of the uniqueness of the solution has been provided yet. In this thesis the calibration problem is formalized by using a unified model that is valid for all central catadioptric omnidirectional cameras. It is also shown that the uniqueness of the problem can be derived from the solution of non-linear equations. However, due to the complexity of the non-linear equations to be solved for the general case, this thesis devises the uniqueness of the solution for the case of a parabolic mirror when the camera's movement is a rotation plus a translation.

Keywords: on-line self-calibration, central catadioptric omnidirectional camera, uncalibrated visual tracking, uniqueness of solution.



MINES ParisTech

ParisTech
INSTITUT DES SCIENCES ET TECHNOLOGIES
PARIS INSTITUTE OF TECHNOLOGY