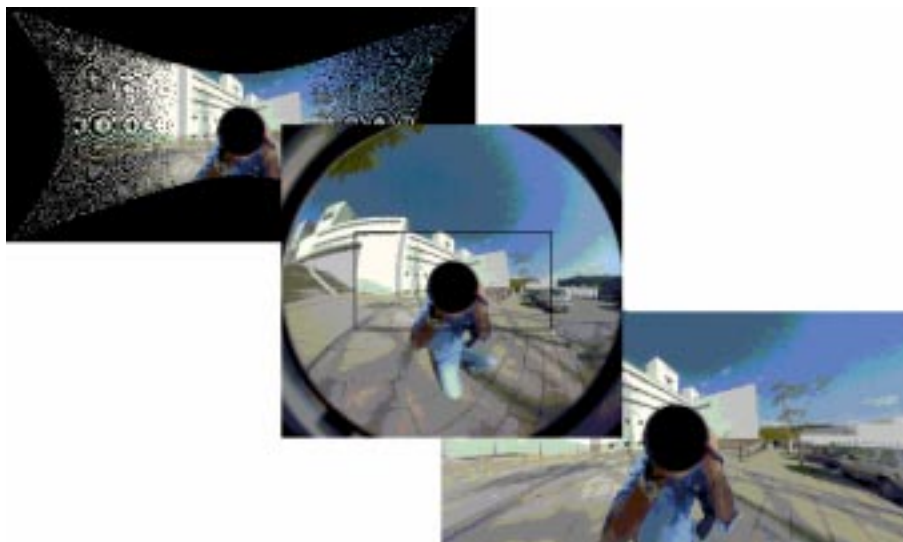


GENERAL CENTRAL PROJECTION SYSTEMS MODELING, CALIBRATION AND VISUAL SERVOING



PHD THESIS
SUBMITTED TO DEPT. OF ELECTRICAL AND COMPUTER ENGINEERING
UNIVERSITY OF COIMBRA

João Pedro de Almeida Barreto

October 2003

© Copyright by João Pedro de Almeida Barreto 2003
All Rights Reserved

Acknowledgements

The first acknowledgement goes to my advisor Prof. Helder Araujo. He was the one who introduced me to computer vision research and helped on my first steps on the field. During these remarkable six years we worked and argued a lot. It was an enjoyable experience even in the bad moments when things were not going so well. This experience would never be possible without his help and advise. Thank you very much for these six years and I hope that we can continue working together in the future.

I also thank the Institute of Systems and Robotics and its president Prof. Traca de Almeida. The ISR has been the institution where I've developed the majority of my research and that mainly supported my expenses and travels to conferences. I also have to acknowledge the Department of Electrical and Computer Engineering as well as the PRODEP program. I would like to mention in particular the Head of the Department, Prof. Luis Sa, for his trust and support to the plan of getting a PRODEP grant.

During these years I've collected a huge list of colleagues that in the meanwhile become my personal friends. The "seniors" Jorge Batista and Paulo Peixoto. Paulo Menezes, who was my advisor during the undergraduate studies (people never forget the first one). And all the others that I met during these years: Johnny and Hugo (my emergency technical advisers); Diogo, Perdigoto, Pi, Nuno, Sergio and Skin (during sometime my undergraduate "slaves"); Paulino, Ana Cristina, Maia, Cortesao ...

The Calouste de Gulbenkian Foundation has partially supported my stay in France in 2002. In France I've to acknowledge the INRIA Rhone-Alpes and the Head of group MOVI, Prof. Radu Horaud, who received me. I also would like to thank all the terrific people that I met. In particular Cristi (my personal optimization teacher), Navneet (my ski mate), Peter and Edmond.

In the United States I thank the GRASP Lab., at the University of Pennsylvania, and Prof. Kostas Daniilidis which made this experience possible.

In a personal level I think that the people that I would like to mention would not fit in this page. Thus I will only say to my family and friends ... thank you very much!

Contents

Acknowledgements	v
1 Introduction	1
1.1 Overview	3
1.2 Contributions	6
2 An Unifying Theory for Central Projection Systems	9
2.1 Central Catadioptric Systems	10
2.2 General Mapping Model	12
2.2.1 Central Catadioptric Image Formation	13
2.2.2 The Unifying Model	16
2.2.3 An Intuitive and “Concrete” Interpretation of the Model	19
2.2.4 About the Non-Linear Function \tilde{h}	20
2.3 Virtual Perspective Images	21
2.4 Closure	22
3 Central Catadioptric Line Projection	25
3.1 The Central Catadioptric Image of Line	25
3.2 The Conic Curve $\bar{\Omega}$	27
3.2.1 Affine Geometry of $\bar{\Omega}$	27
3.2.2 Euclidean Geometry of $\bar{\Omega}$	29
3.3 Projective Properties of the Central Catadioptric Line Image	29
3.3.1 Catadioptric Projection of a Single Line	32
3.3.2 Paracatadioptric Projection of a Single Line	35
3.3.3 Catadioptric Projection of a Pair of Lines	38
3.4 Calibration with Minimal Information	42
3.4.1 Calibrating an Hyperbolic/Elliptical System from Two Lines	42
3.4.2 Full Calibration of a Central Catadioptric System from Three Lines	46
3.5 Closure	49

4	Using Line Images for Calibration Purposes	51
4.1	Working with Conic Curves	51
4.1.1	Intersection of a Line with a Conic	52
4.1.2	Intersection of Two Conic Curves	56
4.2	Fitting a Conic Curve to Image Points	59
4.2.1	Least Square Fitting Based on Algebraic Distances	60
4.2.2	Gradient Weighted Least Squares Fitting	62
4.2.3	Least Squares Fitting Based on Orthogonal Distances	64
4.2.4	Performance Evaluation	66
4.3	Calibration of Central Catadioptric Systems	70
4.3.1	Calibration of an Hyperbolic/Elliptical System	71
4.3.2	Calibration of a General Central Catadioptric System	74
4.3.3	Experiments with Real Images	80
4.4	Closure	82
5	The Paracatadioptric Camera	85
5.1	Paracatadioptric Camera Model	86
5.2	Paracatadioptric Camera Calibration Using Lines	88
5.2.1	Calibration Algorithm	89
5.2.2	Properties of a Set of Paracatadioptric Line Images	91
5.2.3	Estimation of a Set of K Paracatadioptric Line Images	93
5.2.4	Performance Evaluation Using Simulated Images	96
5.2.5	Experimental Results Using Real Images	104
5.3	Direct Least Square Fitting of Paracatadioptric Line Images	106
5.3.1	The Necessary and Sufficient Condition	106
5.3.2	The Algorithm	107
5.3.3	Experiments	109
5.4	Closure	116
6	A General Framework for Selecting the World Coordinate System	117
6.1	World Coordinates for Static Imaging Systems	119
6.1.1	Mapping Points from the 3D Space in the 2D Image Plane	120
6.1.2	Criteria to Select the World Coordinate System	121
6.2	Applying the Framework to Static Imaging Systems	124
6.2.1	Conventional Perspective Camera	124
6.2.2	Central Catadioptric Imaging System	128
6.3	World Coordinates for Imaging Systems with Motion	136
6.3.1	Mapping points from the 3D world into the 2D image plane	137
6.3.2	Criteria to Select the World Coordinate System	138
6.4	Active Tracking of Moving Targets	142

6.4.1	Active Tracking using a Perspective Camera with Translation Motion in the XY Plane	143
6.4.2	Active Tracking using a Perspective Camera with Pan and Tilt Rotation Motion	145
6.4.3	Tracking Applications Using the MDOF and the MVS Robotic platforms	149
6.4.4	Active Tracking Using an Omnidirectional Camera with a Rotational Degree of Freedom Around the Z axis	158
6.5	Closure	162
7	Pose Estimation Using Central Panoramic Imaging	163
7.1	Problem Formulation	164
7.1.1	The Central Panoramic Sensor	164
7.1.2	Measuring the Pose Estimation Error	166
7.1.3	Pose Estimation as a Regulation Control Problem	168
7.2	The Jacobian Matrix for General Central Catadioptric Projection .	169
7.2.1	The Jacobian Matrix J_g	169
7.2.2	Additional Considerations	170
7.3	Model Based Tracking	172
7.4	Closure	172
8	Final Remarks	175
	Bibliography	178

List of Tables

2.1	Equations for the reflective surfaces	11
2.2	Mapping parameters	15
2.3	Multiplying value λ and mapping parameters ξ and φ for the new model	18
3.1	Affine geometry of the conic curve $\bar{\Omega}$	29
3.2	Euclidean geometry of the conic curve $\bar{\Omega}$	30
3.3	Calibrating an hyperbolic/elliptical system from the image of 2 lines	43
3.4	Calibrating a central catadioptric system using 3 lines	47
5.1	Calibrating a paracatadioptric system using K lines ($K \geq 3$)	90
5.2	Calibration results for different number of lines	104
5.3	Recovering angles between pairs of parallel lines	115
6.1	Using a new coordinate system for perspective camera	127
6.2	Using a cartesian coordinate system for perspective camera	127
6.3	Using a new coordinate system for general central catadioptric imaging	133
6.4	Using a spherical coordinate system for active tracking with a pan and tilt perspective camera	146
6.5	Position and velocity command for active tracking with a pan and tilt perspective camera	148
6.6	New coordinate system for active tracking with an omnidirectional camera	160

List of Figures

2.1	The entire class of central catadioptric systems.	10
2.2	Image formation process. Hyperbolic situation	13
2.3	Central catadioptric image formation	14
2.4	New model for central catadioptric image formation	17
2.5	The sphere model for central catadioptric imaging	19
2.6	Generating geometrically correct perspective images	22
2.7	The perspective image of a detail in the scene	23
3.1	The central catadioptric image of a line	26
3.2	The central catadioptric image of a line when $\mathbf{H}_c = \mathbf{I}$	28
3.3	The central catadioptric image of a single line	31
3.4	The central catadioptric image of a pair of lines	39
3.5	Calibrating an hyperbolic/elliptical system from the image of 2 lines	44
3.6	Calibrating a central catadioptric system from the image of 3 lines (Part I)	46
3.7	Calibrating a central catadioptric system from the image of 3 lines (Part II)	48
3.8	Estimating catadioptric line images using standard conic fitting techniques	50
4.1	Intersecting a line with a conic	53
4.2	The tangents to the conic going through a point	54
4.3	Intersection of two conics	55
4.4	Degenerate Configurations	58
4.5	The orthogonal distance between a point and a conic	64
4.6	Conic fitting in the presence of noise	67
4.7	The performance of the conic estimators in the presence of noise .	67
4.8	Conic fitting in the presence of occlusion	69
4.9	The performance of the conic estimators in the presence of occlusion	69
4.10	Calibration of an hyperbolic sensor from 2 line images	72
4.11	Calibration of a general catadioptric sensor from 3 line images . .	75
4.12	Geometric properties of a pair of catadioptric line images	76

4.13	Segmentation of catadioptric line images	80
4.14	Calibration of a general catadioptric system	81
4.15	Convergence for a wrong minima	81
5.1	Model for paracatadioptric image formation	87
5.2	Simulated 480×640 images of three randomly generated lines . .	97
5.3	RMS error in the calibration parameters using the closed form algorithm	98
5.4	Median error, mean error and RMS error	100
5.5	Convergence of the iterative minimization	101
5.6	Median error in the calibration parameters using 3 lines	102
5.7	Median error in the calibration parameters using 3, 5, 7 and 9 lines	103
5.8	Estimating line images for calibration purposes	104
5.9	Perspective obtained by rectifying a paracatadioptric image	105
5.10	Estimating paracatadioptric line image using different methods . .	109
5.11	Comparing the performance of AMS, FF and CATPARB methods ($N = 40, \theta = 80^\circ$)	110
5.12	Comparing the performance of AMS, FF and CATPARB methods with the direct line estimation (DLE) after perspective rectifica- tion ($N = 40, \theta = 80^\circ$)	111
5.13	Characterization of the performance of the CATPARB algorithm. .	113
5.14	Estimating lines in a real paracatadioptric image using only two points	114
5.15	Estimating the angle between pairs of parallel lines	114
6.1	Mapping in central projection systems	119
6.2	Conventional Perspective Camera Case	125
6.3	The sphere model for general central catadioptric image projection	129
6.4	Left: Vertical and horizontal pencils of conic surfaces. Intersec- tion of the lower and upper cone with the unit sphere	131
6.5	Generating a geometrically correct perspective image.	134
6.6	Surveillance application using central panoramic imaging.	135
6.7	Active tracking using a perspective camera with translation motion	143
6.8	Active tracking using a perspective camera mounted on a pan and tilt unit	145
6.9	The Modular Vision System (MVS) at ISR	149
6.10	Distance to the image center after the saccadic motion.	151
6.11	Approximating the global mapping function \mathbf{F} by $\bar{\mathbf{F}}$. Exact (o) and approximated (*) target position.	154

6.12	Statistical analysis of approximation error function \mathbf{E} . Left: Mean of E_x (-) and E_y (- -). Right: Standard deviation of E_x (-) and E_y (- -).	155
6.13	Image velocity field for different situations	157
6.14	Active tracking using an omnidirectional camera	159
7.1	Central catadioptric projection of a rigid body	165
7.2	Iterative pose estimation as a regulation control problem (\mathbf{I} is the $2N \times 2N$ identity matrix).	168
7.3	A tracking sequence. The object translates along axis of camera	172
7.4	A tracking sequence. The object translates in front of camera.	173

Chapter 1

Introduction

Visualization and modeling of large environments is an increasingly attractive proposition. Conventional video cameras provide a limited field of view which can be highly restrictive. Applications that benefit from wide field of view imagery include surveillance, teleconferencing, tele-observation, 3D reconstruction and model acquisition for virtual reality [20, 70, 66, 57]. Enhanced fields of view are also advantageous for visual control of motion applications. In egomotion recovery from video, ambiguities and confusion between translation and rotation may arise whenever the translation direction lies outside the field of view. Panoramic imaging overcomes the problems making the uncertainty of the estimation independent of motion direction [37, 4]. Works on cooperation, obstacle avoidance and self localization of mobile robots also appear in the literature [73, 64, 75, 1]. As shown omnidirectional vision is becoming an increasingly important sub-area in computer vision research [18].

There are two major methods to obtain very wide field of view images. One approach is to build mosaics/panoramas by composing multiple images taken by conventional imaging devices [61]. The alternative method is to use specialized optic-lens arrangements. The approach of combining mirrors with conventional cameras to enhance the sensor field of view is referred to as catadioptric image formation [40]. The tradeoff between the two methods is resolution versus speed of acquisition [45]. In general composing mosaics/panoramas from multiple images provides larger resolutions at the expense of off-line processing. The use of special devices such as catadioptric sensors is simpler and faster enabling the

1. Introduction

capture of dynamic scenes [57].

The single viewpoint constraint is a requirement ensuring that the visual sensor only measures the intensity of light passing through a single point in 3D space (the projection center). Vision systems verifying the single viewpoint constraint are called central projection systems. The well known perspective camera is an example of a central projection system. Central projection systems present interesting geometric properties. A single effective viewpoint is a necessary condition for the generation of geometrically correct perspective images [5], and for the existence of epipolar geometry inherent to the moving sensor and independent of the scene structure [68, 36, 67]. In [5] Baker et al. derives the entire class of catadioptric sensors verifying the single viewpoint constraint. Useful central catadioptric systems can be built by combining a parabolic mirror with an orthographic camera and, an hyperbolic, elliptical or planar mirror with a perspective camera. Despite the nice properties of central projection stated above, non-central catadioptric imaging is also an intense research field [6]. Several sensors have been designed to provide omnidirectional images with specific features. Reflective surfaces capable of providing wide field of view and yet approximate the perspective projection are derived in [41]. A prominent characteristic of most catadioptric sensors is the non-uniform image resolution. Both equi-angular and equi-areal mirror designs aim to improve the resolution uniformity. The equi-angular systems establish a linear relation between angles in the camera and in the mirror [22, 54]. The equi-areal sensors present an area preserving projection from the associated viewing sphere to the image plane [42].

As stated central catadioptric systems combine two important features: a single projection center and a wide field of view. However the mapping between points in the 3D world and points in the image is in general highly non-linear. The single viewpoint constraint assures that geometrically correct perspective images can be generated. Perspective image formation is described by a well known linear model [39]. Thus one strategy to cope with the non-linearities is to generate perspective images from the frames captured by the catadioptric sensor and subsequently process them [5]. This is time consuming and requires an accurate calibration of the catadioptric system. The present work focuses on general central projection systems which include the conventional perspective camera and the catadioptric sensors verifying the single viewpoint constraint. The geometry

1.1. Overview

of central catadioptric image formation is studied in detail and several algorithms are proposed to work directly with the omnidirectional images without warping them.

1.1 Overview

In [34], Geyer et al. introduce an unifying theory for all central catadioptric systems where conventional perspective imaging appears as a particular case. They show that central panoramic projection is isomorphic to a projective mapping from the sphere to a plane with a projection center on the perpendicular to the plane. Chapter 2 introduces a modified version of this unifying model [7, 8]. The mapping between points in the 3D world and points in the catadioptric image plane is split into three steps. 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 referred in [39]). The oriented projective plane is then transformed by a non-linear function \tilde{h} . The last step is a collineation in the plane depending on the mirror parameters, the pose of the camera in relation to the reflective surface and the camera intrinsic parameters. The model obtained is general, intuitive and isolates the non-linear characteristics of general catadioptric image formation.

The established mapping model is used in Chapter 3 to study the geometry of general central catadioptric line projection. Several projective invariant properties of catadioptric line images are derived. These properties are useful for calibration and reconstruction purposes. We show that it is possible to determine the principal point and the image of the absolute conic from three lines in general position [7]. The position of the line at infinity in the catadioptric image plane can be computed from two line images and mirror parameters can be partially recovered without further information [8]. Moreover we prove that if the system is hyperbolic/elliptical and both mirror parameters and camera pose are known then two line images are enough to calibrate the system. The results obtained support the conjecture pointed in [34] that an hyperbolic/elliptical system can be calibrated from a minimum of two lines and a parabolic system requires at least three lines.

In general a line in the scene is mapped into a conic curve in the catadioptric image plane [34]. The calibration algorithms derived in chapter 3 require an accurate estimation of the conic loci where the lines are mapped. However the

1. Introduction

estimation of lines in the catadioptric image plane is hard to accomplish. In general only a small arc of the conic is visible in the image and conventional conic fitting techniques are unable to correctly estimate the curve. Chapter 4 starts by introducing algorithms to compute the intersections between a line and a conic and between two conics. The algorithms are both numerically stable and computationally efficient. The main conic fitting techniques available in the literature are reviewed and evaluated [76, 32]. It is shown that, since they do not cope well with occlusion, they are not suitable to estimate catadioptric lines from image points. The geometric properties derived in chapter 3 are used to constrain the search space and improve the robustness of the conic fitting. This approach leads to non-linear objective functions which must be minimized using iterative gradient descending methods [58, 27]. The performance of the method is evaluated through simulated calibration experiments. It is observed that often the iterative minimization presents convergence problems.

Paracatadioptric sensors combine a parabolic shaped mirror and a camera inducing an orthographic projection. Such a configuration provides a wide field of view while keeping a single effective viewpoint. The paracatadioptric line projection presents specific features. Chapter 5 shows that a set of conic curves corresponds to paracatadioptric line images if, and only if, certain properties are verified [11]. These necessary and sufficient conditions are used to constrain the search space and correctly estimate the curves. The accurate estimation of a minimum of three line images allows the complete calibration of the paracatadioptric camera. If the camera is skewless and the aspect ratio is known then the conic fitting problem is solved naturally by an eigensystem. For the general situation the conic curves are estimated using non-linear optimization. Simulation results are provided to compare the performance of the proposed algorithm with other calibration approaches. Experiments with real images are also presented.

The final part of chapter 5 focuses on line estimation in calibrated paracatadioptric images [9]. The estimation of line images is an important subject for applications such as reconstruction and visual control of motion [12]. However the estimation of the conic curves where lines are mapped is hard to accomplish due to the occlusion problem. A conic curve is the paracatadioptric image of a line if, and only if, the image of the circular points lie on the curve and two certain points are conjugate with respect to the conic. Considering the space of all conic

1.1. Overview

curves, the line images lie in a linear subspace which depends on the system calibration. The paracatadioptric projection of a line can be estimated by fitting a conic in the subspace to the data points. The proposed approach is computationally efficient since the fitting problem can be solved by an eigensystem.

Whenever a central projection system acquires an image, points in 3D space are mapped into points in the 2D image plane. The image formation process represents a transformation from \mathcal{R}^3 to \mathcal{R}^2 , and mathematical models can be used to describe it. Chapter 6 discusses the definition of world coordinate systems that simplify the modeling of general central projection imaging [10]. It is shown that an adequate choice of the world coordinate reference system can be highly advantageous. Such a choice does not imply that new information will be available in the images. Instead the geometric transformations will be represented in a common and more compact framework, while simultaneously enabling newer insights.

The first part of chapter 6 focuses on static central projection systems that include both perspective cameras and catadioptric systems. A systematic approach to select the world reference frame is presented. In particular we derive coordinate systems that satisfy two differential constraints (the "compactness" and the "decoupling" constraints). These coordinate systems have several advantages for the representation of the transformations between the 3D world and the image plane. The second part of the chapter applies the derived mathematical framework to active visual tracking of a target modeled as a moving point in space. In applications of visual control of motion the relationship between motion in the scene and image motion must be established. In the case of active tracking of moving targets these relationships become more complex due to camera motion. Suitable world coordinate reference systems are defined for three distinct situations: perspective camera with planar translation motion, perspective camera with pan and tilt rotation motion, and catadioptric imaging system rotating around an axis going through the effective viewpoint and the camera center. Position and velocity equations relating image motion, camera motion and target 3D motion are derived and discussed. Control laws to perform active tracking of moving targets using visual information are established.

In chapter 6 the moving target is modeled as a point in the scene. However there are several visual servoing applications where the target must be modeled

1. Introduction

has a rigid body with translation and rotation motion. Chapter 7 focuses on iterative pose estimation using central panoramic imaging. The jacobian matrix for general central projection systems is introduced. It is proved that this matrix is a generalization of the well known interaction matrix for conventional perspective cameras [28]. Moreover it is shown that the generalized jacobian has exactly the same singularities as the jacobian of the traditional pinhole camera model. Experiments showing a rigid body being tracked with a catadioptric camera are described.

1.2 Contributions

The most relevant contributions of the present thesis can be summarized as follows

- A modified version of the unifying model for central catadioptric imaging is proposed. The image formation is presented as a three step process with the non-linearities isolated in a function \tilde{h} which maps points between two oriented projective planes.
- The geometry of central catadioptric line images is studied in great detail. Several projective invariant properties are derived. These properties are useful both for calibration and reconstruction tasks.
- It is proved that any central catadioptric system can be fully calibrated using a minimum of three line images in general position.
- It is shown that if the system is hyperbolic/elliptical and both mirror parameters and camera pose are known, then two line images are enough to calibrate the sensor.
- A method to calibrate paracatadioptric sensors using lines is proposed. Geometric constraints are used to accurately estimate the set of lines in the uncalibrated image plane.
- An algorithm to estimate lines in calibrated paracatadioptric images is presented. The proposed method determines the conic locus where the line is mapped using a minimum of two image points. The algorithm is very robust to noise and runs in real time.

1.2. Contributions

- The jacobian matrix for general central catadioptric systems is derived and studied. Experiments on model based tracking and pose estimation of rigid moving bodies are described.
- Any image formation process is a transformation between points in the 3D world and points in the image. There are certain world coordinates systems that are more suitable than others to represent this transformation. We propose a systematic framework to select the world reference frame that simplifies the modeling of a certain system/application.

1. Introduction

Chapter 2

An Unifying Theory for Central Projection Systems

In [34], Geyer et al. introduce an unifying theory for all central catadioptric systems where conventional perspective imaging appears as a particular case. They show that central panoramic projection is isomorphic to a projective mapping from the sphere to a plane with a projection center on the perpendicular to the plane. The present chapter introduces a modified version of this unifying model [7, 8]. The mapping between points in the 3D world and points in the catadioptric image plane is split into three steps. 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 referred in [39]). The oriented projective plane is then transformed by a non-linear function $\tilde{h}(\cdot)$. The last step is a collineation in the plane depending on the mirror parameters, the pose of the camera relative to the reflective surface and the camera intrinsic parameters. The model obtained is general, intuitive and isolates the non-linear characteristics of general catadioptric image formation.

The chapter starts by presenting the entire class of central catadioptric systems derived in [5]. The image formation process is investigated and the general mapping model is derived. The conventional perspective camera is a particular case of the presented unifying theory.

2. An Unifying Theory for Central Projection Systems

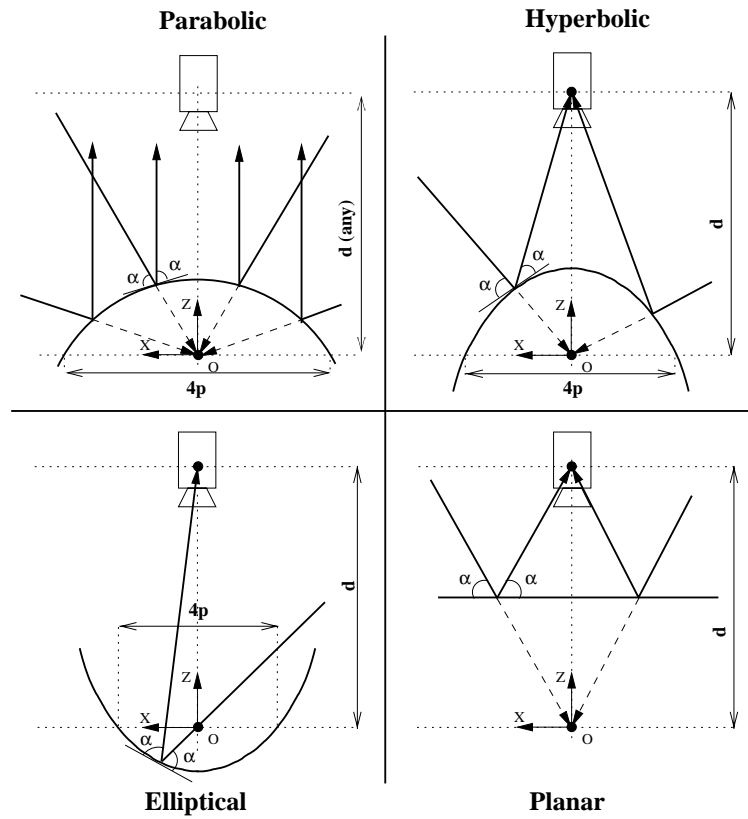


Figure 2.1: The entire class of central catadioptric systems.

2.1 Central Catadioptric Systems

Central catadioptric systems provide a wide field of view while keeping an unique projection center. The wide field of view is achieved by specialized optic-lens arrangements. The single viewpoint constraint is verified by a careful selection and assembly of mirror and imaging device. The final vision system must only measure the intensity of light passing through a single point in 3D space (the projection center).

Consider a reflective surface S , a generic point P lying on S and a plane Π going through point P and tangent to the mirror surface. Any ray of light r_i incident on P is reflected in a ray of light r_r . According to the electromagnetic theory the angle between r_i and plane Π (the incident angle) must be equal to the angle that r_r makes with the same plane (the reflection angle).

2.1. Central Catadioptric Systems

Parabolic	$\sqrt{x^2 + y^2 + z^2} = 2p - z$
Hyperbolic	$\frac{(z-\frac{d}{2})^2}{a^2} - \frac{x^2}{b^2} - \frac{y^2}{b^2} = 1$
Elliptical	$\frac{(z-\frac{d}{2})^2}{a^2} + \frac{x^2}{b^2} + \frac{y^2}{b^2} = 1$
Planar	$z = \frac{d}{2}$

Table 2.1: Equations for the reflective surfaces

Fig.2.1 shows the four types of central catadioptric systems [5]. A scheme of a parabolic mirror with latus rectum $4p$ is exhibited on the top-left corner. Tab.2.1 provides the equation of the 3D surface assuming the origin O of the coordinate system coincident with the focus of the paraboloid and the z -axis aligned with the mirror axis. It is a well known result that any incident ray of light going through the focus of the paraboloid is reflected into a ray parallel to the surface axis. If the camera steering the mirror is orthographic and the image plane is orthogonal to the z -axis then it only captures the rays parallel to the mirror axis. These rays result from the reflection of the light going through the focus of the mirror. The vision system has a single effective viewpoint which is the focus of the parabolic mirror.

The mirror in the top-right corner of Fig. 2.1 has an hyperbolic shape with a latus rectum $4p$ and a distance between foci d . The corresponding hyperboloid equation is provided in Tab. 2.1 where $a = \frac{1}{2}(\sqrt{d^2 + 4p^2} - 2p)$ and $b = \sqrt{p(\sqrt{d^2 + 4p^2} - 2p)}$. Once again the origin O of the coordinate system is coincident with the inner focus and the z -axis is aligned with the surface symmetry axis. From the reflection law it comes that any light ray going through the inner focus is reflected in another ray passing through the outer focus. A conventional perspective camera positioned in the outer focus only measures the intensity of light going through O . A perspective camera steering and hyperbolic mirror such that its projection center is coincident with the outer focus is a central catadioptric system. The effective viewpoint is the inner focus of the reflective surface. In a similar way a central catadioptric system can be built by combining a perspective camera with an elliptical mirror (see bottom-left corner of Fig. 2.1). The equation of the ellipsoid is provided in Tab. 2.1 where $4p$ is the latus rectum, d is the

2. An Unifying Theory for Central Projection Systems

distance between foci, $a = \frac{1}{2}(\sqrt{d^2 + 4p^2} + 2p)$ and $b = \sqrt{p(\sqrt{d^2 + 4p^2} + 2p)}$.

A catadioptric system made up of a perspective camera steering a planar mirror also verifies the single view point constraint (bottom-right corner of Fig. 2.1). The effective projection center is behind the mirror in the perpendicular line passing through camera center. Its distance to the camera center is twice the distance between the planar mirror and the camera.

For the parabolic system the distance between the camera and the mirror is not constrained. The single viewpoint constraint is verified whenever the camera is orthographic and optical axis is aligned with the axis of the paraboloid. For the hyperbolic and elliptical systems the center of the perspective camera must be coincident with the outer focus of the reflective surface. For the planar situation changes in camera position imply changes in the effective viewpoint. However in these three cases the alignment between the camera and the mirror is not constrained. After rotating the camera around an axis going through its center the catadioptric systems still verifies the single viewpoint constraint and the projection center position is kept.

As a final remark notice that the field of view is enhanced by using convex reflective surfaces like paraboloid and hyperboloid. Elliptical mirrors do not have the same practical application because they cause a decrease in the field of view. In the elliptical systems the field of view is traded by resolution which causes a zoom effect in the images. However the mirror introduces non-linearities in the imaging formation which is a disadvantage if compared with conventional cameras equipped with zoom lens. It will be shown that the planar system is in many aspects a degenerate case of the central catadioptric projection. The geometry of planar catadioptric images is equivalent to conventional perspective imaging. Nevertheless planar mirrors are currently used to build devices like the Nalwa pyramid which is a high resolution omnidirectional system with a single projection center [52].

2.2 General Mapping Model

The present section studies the image formation process in central catadioptric systems and derives an unifying mapping model.

2.2. General Mapping Model

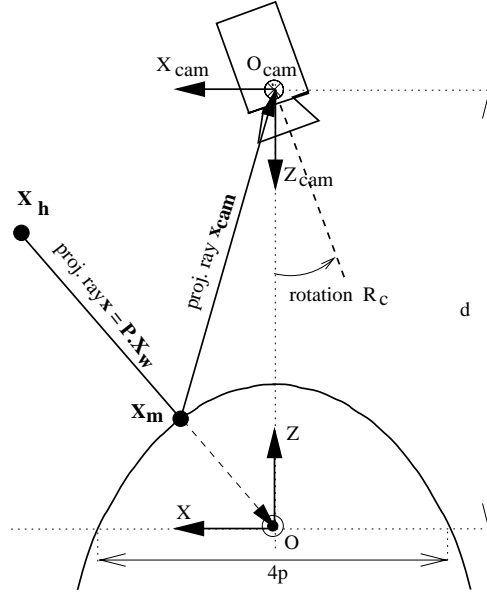


Figure 2.2: Image formation process. Hyperbolic situation

2.2.1 Central Catadioptric Image Formation

Consider the coordinate systems \mathcal{R} and \mathcal{R}_{cam} depicted in Fig.2.2. \mathcal{R} is the coordinate system attached to the reflective surface. The z -axis is aligned with the mirror axis and the origin O is coincident with the effective viewpoint of the central projection system. For parabolic, hyperbolic and elliptical systems the origin O is the inner focus of the reflective surface. The transformation matrix from \mathcal{R} to \mathcal{R}_{cam} is provided in equation 2.1. The z -axis of \mathcal{R}_{cam} is still aligned with the mirror axis and the distance between the origins O and O_{cam} is d . For both hyperbolic and elliptical systems d corresponds to the distance between foci and the central catadioptric images are acquired by a perspective camera with projection center in O_{cam} . The planar system has a perspective camera positioned in O_{cam} in a similar manner. For the parabolic situation the distance d is unconstrained and an orthographic camera with optical axis parallel to the z -axis of \mathcal{R}_{cam} is used.

2. An Unifying Theory for Central Projection Systems

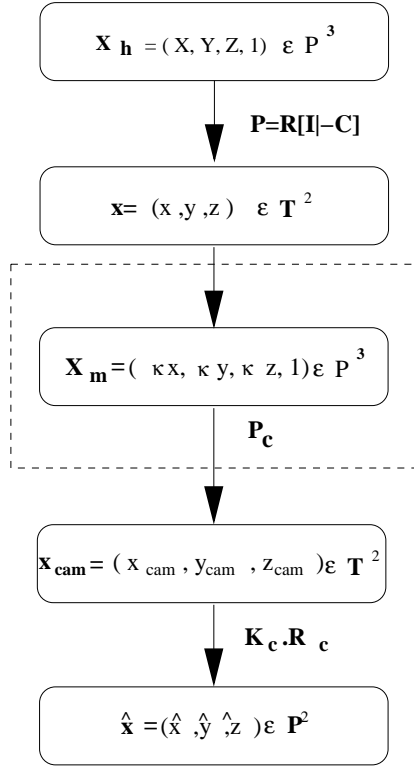


Figure 2.3: Central catadioptric image formation

$$\mathbf{T}_{\text{cam}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & d \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.1)$$

Fig.2.3 is a step-by-step scheme of the mapping performed by a general central catadioptric system. Consider a generic scene point visible by the catadioptric system as depicted in Fig. 2.2. \mathbf{X}_h is the corresponding vector of homogeneous coordinates in the world reference frame. Each visible point can be associated to a projective ray \mathbf{x} joining the point with the effective viewpoint of the central projection system. $\mathbf{x} = \mathbf{P}\mathbf{X}_h$ where $\mathbf{P} = \mathbf{R}[\mathbf{I} - \mathbf{C}]$ is a 3×4 matrix transforming points in the world reference frame in projective rays in the coordinate system \mathfrak{R} attached to the mirror (\mathbf{C} represents the world origin coordinates in the mirror reference frame, \mathbf{R} is the rotation matrix between the two coordinate systems and

2.2. General Mapping Model

	κ	\mathbf{P}_c	\mathbf{R}_c
Parabolic	$\frac{2p}{z + \sqrt{x^2 + y^2 + z^2}}$	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$	\mathbf{I}
Hyperbolic	$\frac{2dp}{z(\sqrt{d^2 + 4p^2 + 2p}) + d\sqrt{x^2 + y^2 + z^2}}$	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & d \end{bmatrix}$	any
Elliptical	$-\frac{2dp}{z(\sqrt{d^2 + 4p^2 - 2p}) + d\sqrt{x^2 + y^2 + z^2}}$	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & d \end{bmatrix}$	any
Planar	$\frac{d}{2z}$	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & d \end{bmatrix}$	any

Table 2.2: Mapping parameters

\mathbf{I} is a 3×3 identity matrix). We can think of the projective rays \mathbf{x} as points in an oriented projective plane \mathbf{T}^2 . Notice that in standard projective geometry, given a projective point \mathbf{x} , $\lambda\mathbf{x}$ represents the same point whenever $\lambda \neq 0$. In an oriented projective plane this is only true if $\lambda > 0$ [47, 65]. This is important when modeling panoramic vision sensors where diametrically opposite points relative to the projection center can be simultaneously imaged.

The projective ray \mathbf{x} intersects the mirror surface on point \mathbf{X}_m (see Fig. 2.2). The intersection point can be computed by scaling the projective ray \mathbf{x} by a κ value, such that $\kappa\mathbf{x}$ verifies the mirror equation. Thus, replacing (x, y, z) by $(\kappa x, \kappa y, \kappa z)$ in the surface equations of Tab. 2.1 and solving them in order to κ , the scaling values for each type of central catadioptric system are obtained (see Tab. 2.2). The intersection point in mirror homogeneous coordinates is $\mathbf{X}_m = (\kappa\mathbf{x}, 1)^t$. Notice that κ depends both on \mathbf{x} and on the mirror surface.

For the hyperbolic, elliptical and planar catadioptric systems the imaging device is a conventional perspective camera with center on the origin \mathbf{O}_{cam} of the coordinate system $\mathfrak{R}_{\text{cam}}$. Fig. 2.2 shows that to each intersection point \mathbf{X}_m corresponds a projective ray \mathbf{x}_{cam} going through the camera center. Consider the pin-hole camera model $\mathbf{P}_p = [\mathbf{I}|\mathbf{0}]$ and the coordinate transformation matrix \mathbf{T}_{cam} of equation 2.1. The projective ray coordinates \mathbf{x}_{cam} in $\mathfrak{R}_{\text{cam}}$ are given by $\mathbf{x}_{\text{cam}} = \mathbf{P}_c\mathbf{X}_m$ with $\mathbf{P}_c = \mathbf{P}_p\mathbf{T}_{\text{cam}}$ (see Tab. 2.2). For the parabolic system an orthographic camera must be used instead of a perspective camera. The camera center is at infinity and the pin-hole model \mathbf{P}_p must be replaced by the orthographic camera model \mathbf{P}_o . In a similar manner the projective ray is

2. An Unifying Theory for Central Projection Systems

$\mathbf{x}_{\text{cam}} = \mathbf{P}_c \mathbf{X}_m$ with $\mathbf{P}_c = \mathbf{P}_o \mathbf{T}_{\text{cam}}$. Notice in Tab. 2.2 that the resultant matrix \mathbf{P}_c no longer depends on the distance d .

The world point \mathbf{X}_h is imaged in $\hat{\mathbf{x}} = \mathbf{K}_c \mathbf{R}_c \mathbf{x}_{\text{cam}}$ at the central catadioptric image plane. \mathbf{K}_c is the matrix of camera intrinsic parameters and \mathbf{R}_c is a 3×3 rotation matrix. As stated the hyperbolic, elliptical and planar systems have a single viewpoint whenever the perspective camera center is in the origin \mathbf{O}_{cam} of $\mathfrak{R}_{\text{cam}}$. However the camera pose is not constrained and the optical axis is not necessarily aligned with the mirror axis. Matrix \mathbf{R}_c models the possible rotation of the camera with respect to the coordinate system $\mathfrak{R}_{\text{cam}}$ (see Fig. 2.2). For the parabolic system the camera is orthographic and the image plane must be orthogonal to the mirror symmetry axis. The requirement of a single effective viewpoint constraints the camera pose which can only rotate around the z-axis of $\mathfrak{R}_{\text{cam}}$. It is assumed without loss of generality that $\mathbf{R}_c = \mathbf{I}$ for the central parabolic sensor (see Tab. 2.2).

2.2.2 The Unifying Model

The mapping scheme of Fig. 2.3 with parameters on Tab. 2.2 is general, covering all central catadioptric systems, in all possible configurations. Visible points in the scene \mathbf{X}_h are mapped into projective rays/points \mathbf{x} in the catadioptric system reference frame centered in the effective viewpoint. The transformation is linear being described by a 3×4 matrix \mathbf{P} . To each oriented projective ray/point \mathbf{x} , corresponds a projective ray/point \mathbf{x}_{cam} in coordinate system $\mathfrak{R}_{\text{cam}}$ (see Fig. 2.2). The relationship between projective point $\hat{\mathbf{x}}$ measured in the catadioptric image plane and \mathbf{x}_{cam} is established by a collineation depending on camera orientation and intrinsic parameters.

For a conventional projective camera the mapping between points in the world and points in the image is linear if homogeneous coordinates are assumed [39]. In the model established for general central catadioptric imaging all the transformations are linear with the exception of the mapping of \mathbf{x} into \mathbf{x}_{cam} . As mentioned before \mathbf{x} and \mathbf{x}_{cam} are oriented projective rays/points that must intersect on the mirror surface. The relationship between these two points is established by $\mathbf{x}_{\text{cam}} = \mathbf{P}_c(\kappa \mathbf{x}, 1)^t$. Matrix \mathbf{P}_c depends on the mirror parameters and on the type of imaging device. The multiplying parameter κ is a function of the mirror

2.2. General Mapping Model

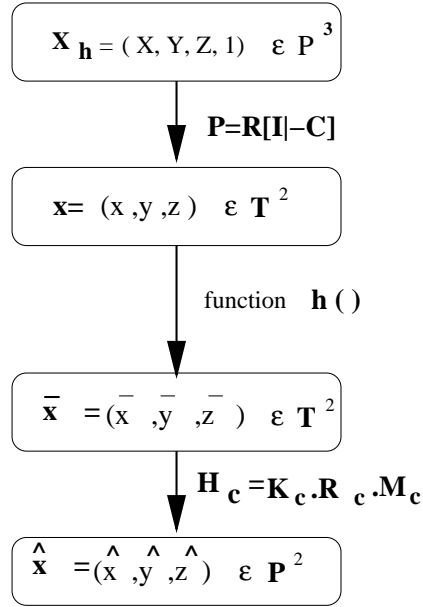


Figure 2.4: New model for central catadioptric image formation

parameters and of point coordinates \mathbf{x} (see Tab. 2.2). The transformation of \mathbf{x} into \mathbf{x}_{cam} can be seen as a non-linear mapping between two oriented projective planes. Since \mathbf{x}_{cam} is an oriented projective point, then $\lambda \mathbf{x}_{\text{cam}}$ represents the same point whenever $\lambda > 0$ [65, 47]. The relationship between \mathbf{x} and \mathbf{x}_{cam} can be written as

$$\mathbf{x}_{\text{cam}} = \lambda \left(\mathbf{P}_c \cdot \begin{bmatrix} \kappa \mathbf{x} \\ 1 \end{bmatrix} \right) \quad (2.2)$$

Replace in equation 2.2 \mathbf{P}_c , κ and λ by the values provided on tables 2.2 and 2.3. The values for the parameters are selected according to the type of central catadioptric system. Notice that λ is always greater than zero. After some algebraic manipulation you will verify that the transformation of \mathbf{x} into \mathbf{x}_{cam} can be written in the form of equation 2.3. Matrix \mathbf{M}_c and function \tilde{h} are respectively provided by equations 2.4 and 2.5. Parameters ξ and φ appear on Tab. 2.3. Matrix \mathbf{M}_c only depends on mirror type and shape.

$$\mathbf{x}_{\text{cam}} = \mathbf{M}_c \cdot \tilde{h}(\mathbf{x}) \quad (2.3)$$

2. An Unifying Theory for Central Projection Systems

	λ	ξ	φ
Parabolic	$\frac{z}{\sqrt{x^2+y^2+z^2}} + 1$	1	$1 + 2p$
Hyperbolic	$\frac{z(\sqrt{d^2+4p^2}+2p)}{d\sqrt{x^2+y^2+z^2}\sqrt{d^2+4p^2}} + \frac{1}{\sqrt{d^2+4p^2}}$	$\frac{d}{\sqrt{d^2+4p^2}}$	$\frac{d+2p}{\sqrt{d^2+4p^2}}$
Elliptical	$\frac{z(\sqrt{d^2+4p^2}-2p)}{d\sqrt{x^2+y^2+z^2}\sqrt{d^2+4p^2}} + \frac{1}{\sqrt{d^2+4p^2}}$	$\frac{d}{\sqrt{d^2+4p^2}}$	$\frac{d-2p}{\sqrt{d^2+4p^2}}$
Planar	$\frac{2z}{d\sqrt{x^2+y^2+z^2}}$	0	1

Table 2.3: Multiplying value λ and mapping parameters ξ and φ for the new model

$$\mathbf{M}_c = \begin{bmatrix} \varphi - \xi & 0 & 0 \\ 0 & \xi - \varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2.4)$$

$$\tilde{h}(\mathbf{x}) = \begin{bmatrix} \frac{x}{\sqrt{x^2+y^2+z^2}} \\ \frac{y}{\sqrt{x^2+y^2+z^2}} \\ \frac{z}{\sqrt{x^2+y^2+z^2}} + \xi \end{bmatrix} \quad (2.5)$$

The scheme of Fig. 2.4 is obtained by rearranging the one of Fig. 2.3. The mapping between points in the world \mathbf{X}_h and projective image points $\hat{\mathbf{x}}$ is given by equation 2.6. The mapping model is general for all central catadioptric systems. Depending on the type of system the parameters ξ and φ of function \tilde{h} and matrix \mathbf{M}_c change according to Tab. 2.3. Points \mathbf{X}_h in projective 3D space are transformed in points \mathbf{x} in the oriented projective plane with origin in the effective viewpoint ($\mathbf{x} = \mathbf{P}\mathbf{X}_h$). Points \mathbf{x} are mapped into points $\bar{\mathbf{x}}$ in a second oriented projective plane. The correspondence function $\bar{\mathbf{x}} = \tilde{h}(\mathbf{x})$ is non-linear. Projective points $\hat{\mathbf{x}}$ in catadioptric image plane are obtained after a collineation \mathbf{H}_c ($\hat{\mathbf{x}} = \mathbf{H}_c\bar{\mathbf{x}}$ with \mathbf{H}_c provided by equation 2.7).

$$\hat{\mathbf{x}} = \mathbf{H}_c \underbrace{\tilde{h}(\mathbf{P}\mathbf{X}_h)}_{\bar{\mathbf{x}}} \quad (2.6)$$

$$\mathbf{H}_c = \mathbf{K}_c \mathbf{R}_c \mathbf{M}_c \quad (2.7)$$

2.2. General Mapping Model

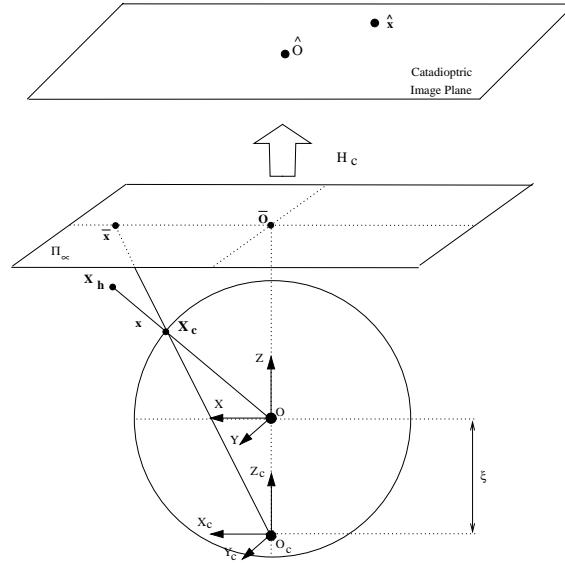


Figure 2.5: The sphere model for central catadioptric imaging

2.2.3 An Intuitive and “Concrete” Interpretation of the Model

Fig. 2.4 is a schematic of the proposed unifying model for central catadioptric image formation. The mapping model is made up of three steps. The first step is the linear transformation $\mathbf{P} = \mathbf{R}[\mathbf{I} - \mathbf{C}]$ of 3D world coordinates into 2D sensor coordinates. If nothing is said it will be assumed without loss of generality, that world and sensor coordinates are the same. Thus, since $\mathbf{R} = \mathbf{I}$ (no relative rotation) and $\mathbf{C} = \mathbf{0}$ (coincident origins), arises that $\mathbf{P} = [\mathbf{I}|\mathbf{0}]$. The second step is the non-linear mapping \tilde{h} between two oriented projective planes. The last step is the projective transformation \mathbf{H}_c depending on the camera calibration, the mirror parameters and their relative pose (equation 2.7). The relationship between points in the catadioptric image plane $\hat{\mathbf{x}}$ and points $\bar{\mathbf{x}}$ is linear ($\hat{\mathbf{x}} = \mathbf{H}_c\bar{\mathbf{x}}$).

The proposed model isolates the non-linearities of the mapping in a single function \tilde{h} . Function \tilde{h} transforms oriented projective points \mathbf{x} in sensor coordinates into points $\bar{\mathbf{x}}$. This non-linear transformation, presented in equation 2.5, has an intuitive “concrete” interpretation. Consider the coordinate system \mathfrak{R} , with origin \mathbf{O} in the effective viewpoint, and an unit sphere centered in \mathbf{O} (see Fig. 2.5). 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

2. An Unifying Theory for Central Projection Systems

unit sphere in a single point \mathbf{X}_c . Consider a point \mathbf{O}_c with coordinates $(0, 0, -\xi)^t$ in \mathfrak{R} . To each \mathbf{x} corresponds an oriented projective ray $\bar{\mathbf{x}}$ joining \mathbf{O}_c with the intersection point \mathbf{X}_c in the sphere surface. The non-linear mapping \hbar corresponds to projecting the scene in the unity sphere surface and then re-projecting the points on the sphere into a plane from a novel projection center \mathbf{O}_c . Points in catadioptric image plane $\hat{\mathbf{x}}$ are obtained after a collineation \mathbf{H}_c of 2D projective points $\bar{\mathbf{x}}$.

The novel projection center $\mathbf{O}_c = (0, 0, -\xi)^t$ only depends on mirror parameters (see Table.2.3). For a parabolic mirror $\xi = 1$ and \mathbf{O}_c belongs to the sphere surface. The re-projection is a stereographic projection. For hyperbolic and elliptical case \mathbf{O}_c is inside the sphere in 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} . The mapping of equation 2.6 becomes linear and the model is equivalent to the conventional model for projective cameras.

2.2.4 About the Non-Linear Function \hbar

The mathematical expression of function \hbar is provided by equation 2.6. Notice that $\hbar(\lambda\mathbf{x}) = \lambda\hbar(\mathbf{x})$ whenever $\lambda > 0$. Thus \hbar is a positive homogeneous function. To each oriented projective ray \mathbf{x} corresponds one, and only one, oriented projective ray $\bar{\mathbf{x}}$ such that $\bar{\mathbf{x}} = \hbar(\mathbf{x})$. Since function \hbar is injective then it has an inverse \hbar^{-1} .

The inverse function \hbar^{-1} maps oriented projective points $\bar{\mathbf{x}}$ into oriented projective points \mathbf{x} ($\mathbf{x} = \hbar^{-1}(\bar{\mathbf{x}})$). Consider the system of coordinates \mathfrak{R}_c in Fig. 2.5. To each projective ray $\bar{\mathbf{x}}$ corresponds a point \mathbf{X}_c lying on the surface of the sphere. Assuming $\bar{\mathbf{x}} = (\bar{x}, \bar{y}, \bar{z})^t$ there is a positive multiplying parameter λ_c such that $\mathbf{X}_c = (\lambda_c\bar{x}, \lambda_c\bar{y}, \lambda_c\bar{z})^t$. Changing to sensor coordinates \mathfrak{R} comes that $\mathbf{X}_c = (\lambda_c\bar{x}, \lambda_c\bar{y}, \lambda_c\bar{z} - \xi)^t$. Since \mathbf{X}_c lies in the unitary sphere centered in the origin \mathbf{O} then $\lambda_c^2\bar{x}^2 + \lambda_c^2\bar{y}^2 + (\lambda_c\bar{z} - \xi)^2 = 1$. Solving the equation in order to λ_c and choosing the positive solution yields

$$\lambda_c = \frac{\bar{z}\xi + \sqrt{\bar{z}^2 + (1 - \xi^2)(\bar{x}^2 + \bar{y}^2)}}{\bar{x}^2 + \bar{y}^2 + \bar{z}^2} \quad (2.8)$$

The 3D point \mathbf{X}_c lies on the oriented projective ray \mathbf{x} . Thus the projective coordinates of \mathbf{x} in the sensor reference frame are $\mathbf{x} = (\lambda_c\bar{x}, \lambda_c\bar{y}, \lambda_c\bar{z} - \xi)^t$. The

2.3. Virtual Perspective Images

mathematical expression of the inverse function h^{-1} is

$$h^{-1}(\bar{\mathbf{x}}) = \begin{bmatrix} \frac{\bar{z}\xi + \sqrt{\bar{z}^2 + (1-\xi^2)(\bar{x}^2 + \bar{y}^2)}}{\bar{x}^2 + \bar{y}^2 + \bar{z}^2} \bar{x} \\ \frac{\bar{z}\xi + \sqrt{\bar{z}^2 + (1-\xi^2)(\bar{x}^2 + \bar{y}^2)}}{\bar{x}^2 + \bar{y}^2 + \bar{z}^2} \bar{y} \\ \frac{\bar{z}\xi + \sqrt{\bar{z}^2 + (1-\xi^2)(\bar{x}^2 + \bar{y}^2)}}{\bar{x}^2 + \bar{y}^2 + \bar{z}^2} \bar{z} - \xi \end{bmatrix} \quad (2.9)$$

2.3 Virtual Perspective Images

As referred in [5] a single projection center is a necessary condition for the generation of geometrically correct perspective images. This section presents two different approaches to artificially generate perspective images from central catadioptric images. It is assumed that the vision system is calibrated and matrix \mathbf{H}_c is known (equation 2.7). The projection center of the virtual perspective camera must be coincident with the effective viewpoint \mathbf{O} (see Fig. 2.5). The matrix of intrinsic parameters is \mathbf{K}_v and the rotation matrix between the reference frame attached to the virtual camera and the sensor coordinate system is \mathbf{R}_v . Both matrices are defined in advance by the user.

The mapping model derived in the previous sections is schematized in Fig. 2.4. It has been shown that to each point $\hat{\mathbf{x}}$ in the catadioptric image plane corresponds one, and only one, projective ray \mathbf{x} . The mapping function is $\mathbf{x} = h(\mathbf{H}_c^{-1}\hat{\mathbf{x}})$. Assume that \mathbf{x}_v is a perspective image point, provided in projective coordinates in the reference frame \mathfrak{R}_v attached to the virtual camera. The mathematical relation between points \mathbf{x}_v and \mathbf{x} is $\mathbf{x}_v = \mathbf{K}_v\mathbf{R}_v\mathbf{x}$. Replacing \mathbf{x} yields the result of equation 2.10 which maps points $\hat{\mathbf{x}}$ in the catadioptric image into points \mathbf{x}_v in the virtual perspective image.

$$\mathbf{x}_v = \mathbf{K}_v\mathbf{R}_v h^{-1}(\mathbf{H}_c^{-1}\hat{\mathbf{x}}) \quad (2.10)$$

Fig. 2.6 shows an omnidirectional image with dimension 2272×1704 acquired by a paracatadioptric camera [35]. A virtual perspective image is exhibited on the top right corner. The image is generated using the result of equation 2.10 with $\mathbf{R}_v = \mathbf{I}$. Points inside the rectangle marked in the omnidirectional frame are mapped into points in the virtual perspective image. Notice that the scene is not uniformly sampled by the catadioptric sensor. In Fig. 2.7 the same strategy is

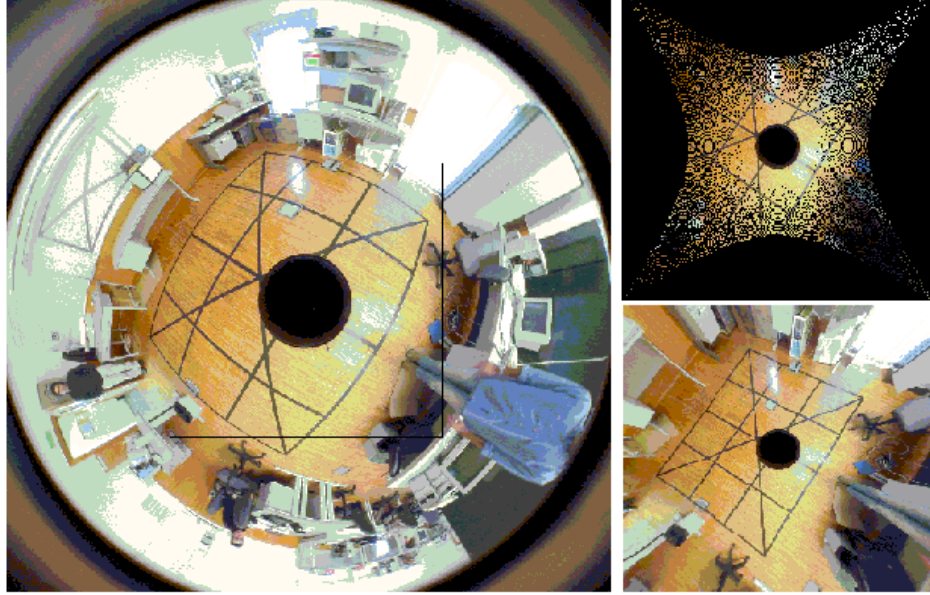


Figure 2.6: Generating geometrically correct perspective images

used to generate a perspective image of the tile panel in the wall. In this situation $\mathbf{R}_v \neq \mathbf{I}$.

An alternative approach can be used to generate perspective images more suitable for visualization purposes. Equation 2.11 is the inverse of the mapping provided in equation 2.10. All points \mathbf{x}_v in the perspective image plane are transformed into points in the catadioptric image using the derived relation. The brightness (or color) in the perspective image points is then computed using bilinear interpolation. The results can be observed at the bottom right corners of Fig. 2.6 and 2.7.

$$\hat{\mathbf{x}} = \mathbf{H}_c \mathbf{h}(\mathbf{R}_v^{-1} \mathbf{K}_v^{-1} \mathbf{x}_v) \quad (2.11)$$

2.4 Closure

In this chapter we have reviewed the entire class of central catadioptric systems. Central catadioptric systems combine two useful features: a wide field of view and a single projection center [5]. Central catadioptric image formation is isomorphic

2.4. Closure

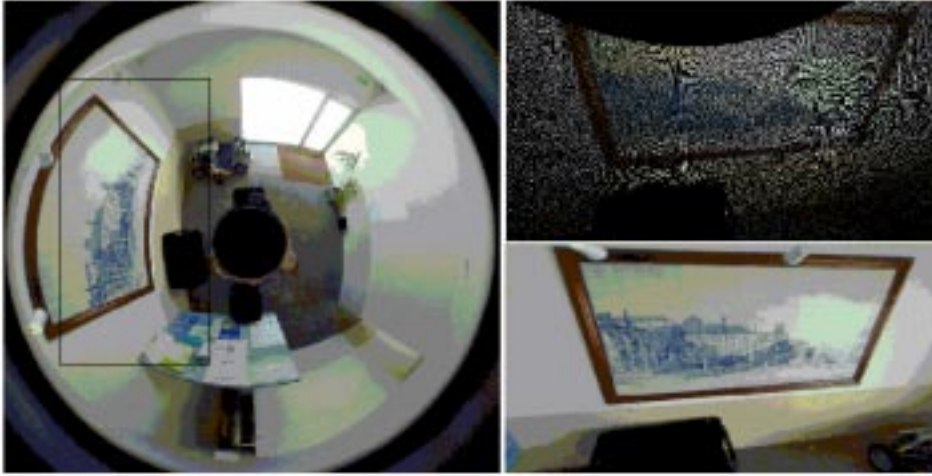


Figure 2.7: The perspective image of a detail in the scene

to a projective mapping from the sphere to a plane with projection center in the perpendicular to the plane [34]. This chapter shows that the image formation can be modeled as a three step process. A conventional 3×4 projection matrix \mathbf{P} maps the 3D scene in an oriented projective plane. The oriented projective plane is transformed by a non-linear function \tilde{h} . The resulting oriented projective plane is mapped in the final catadioptric image by a collineation which depends on the camera intrinsic parameters and on the relative pose between the imaging sensor and the reflective surface. The proposed model is general, intuitive and isolates the non linearities in an injective function \tilde{h} .

2. An Unifying Theory for Central Projection Systems

Chapter 3

Central Catadioptric Line Projection

The mapping model derived on the previous chapter is made up of three steps: a linear transformation \mathbf{P} of 3D world coordinates into 2D sensor coordinates, a non-linear mapping \tilde{h} between two oriented projective planes, and a final 2D collineation \mathbf{H}_c (see Fig. 2.4). The present chapter studies the line projection for central catadioptric systems. It is shown that in general a line is imaged into a conic curve. Collineation \mathbf{H}_c is ignored in section 3.2. By assuming $\mathbf{H}_c = \mathbf{I}$ we focus on the effects of the non-linear mapping \tilde{h} in the catadioptric line projection. Both affine and euclidean geometry of the resultant conic curve are studied. However in general $\mathbf{H}_c \neq \mathbf{I}$. Section 3.3 derives projective invariant properties of central catadioptric line projection.

3.1 The Central Catadioptric Image of Line

Consider a line in space lying in a plane $\Pi = (n_x, n_y, n_z, 0)^t$ which contains the effective viewpoint \mathbf{O} (Fig. 3.1). According to the first step of the mapping model of Fig. 2.4, the 3D line projects in $\Pi = \mathbf{P}^t \mathbf{n}$. Since it is assumed that the world reference frame and the sensor system of coordinates are the same, then $\mathbf{P} = [\mathbf{I}|\mathbf{0}]$ and $\mathbf{n} = (n_x, n_y, n_z)^t$. Thus the world points \mathbf{X}_w lying on the original line are mapped into points \mathbf{x} in the oriented projective plane such that $\mathbf{n}^t \cdot \mathbf{x} = 0$.

The non-linear function \tilde{h} establishes the relationship between points \mathbf{x} and

3. Central Catadioptric Line Projection

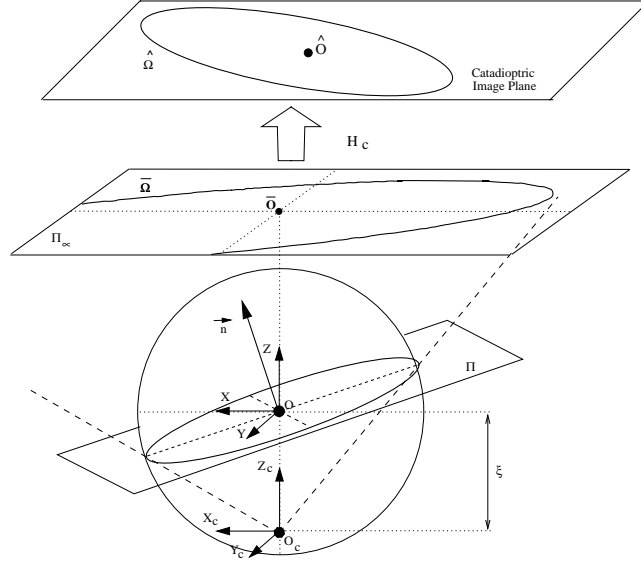


Figure 3.1: The central catadioptric image of a line

$\bar{\mathbf{x}}$. Consider the inverse function \bar{h}^{-1} provided in equation 2.9. If $\mathbf{n}^t \cdot \mathbf{x} = 0$ and $\mathbf{x} = \bar{h}^{-1}(\bar{\mathbf{x}})$ then $\mathbf{n}^t \cdot \bar{h}^{-1}(\bar{\mathbf{x}}) = 0$. After some algebraic manipulation the equality can be written in the form $\bar{\mathbf{x}}^t \bar{\mathbf{\Omega}} \bar{\mathbf{x}} = 0$ with $\bar{\mathbf{\Omega}}$ given by equation 3.1. The non-linear mapping \bar{h} between two oriented projective planes transforms the line \mathbf{n} into a conic curve $\bar{\mathbf{\Omega}}$.

$$\bar{\mathbf{\Omega}} = \begin{bmatrix} n_x^2(1 - \xi^2) - n_z^2\xi^2 & n_x n_y(1 - \xi^2) & n_x n_z \\ n_x n_y(1 - \xi^2) & n_y^2(1 - \xi^2) - n_z^2\xi^2 & n_y n_z \\ n_x n_z & n_y n_z & n_z^2 \end{bmatrix} \quad (3.1)$$

Points in the catadioptric image plane $\hat{\mathbf{x}}$ are linearly related with points $\bar{\mathbf{x}}$ through a collineation \mathbf{H}_c (equation 2.6). The projective transformation of a conic curve is always a conic curve [59, 39, 62]. Thus conic $\bar{\mathbf{\Omega}}$ is mapped in the catadioptric image plane into a conic $\hat{\mathbf{\Omega}}$ (equation 3.2). In general we may conclude that a line in the scene is projected into a conic curve $\hat{\mathbf{\Omega}}$ in the catadioptric image plane [34].

$$\hat{\mathbf{\Omega}} = \mathbf{H}_c^{-t} \bar{\mathbf{\Omega}} \mathbf{H}_c^{-1} \quad (3.2)$$

Fig.3.1 depicts central catadioptric line projection using the sphere model. The

3.2. The Conic Curve $\bar{\Omega}$

world line in space is projected into a great circle on the sphere surface. This great circle is the curve of intersection of plane Π , containing both the line and the projection center \mathbf{O} , and the unit sphere. The projective rays $\bar{\mathbf{x}}$, joining \mathbf{O}_c to points in the great circle, form a central cone surface [63]. The central cone, with vertex in \mathbf{O}_c , projects into the conic $\bar{\Omega}$ in the canonical image plane. Notice that we can always think of a conic $\bar{\Omega}$ in the projective plane as a central cone of projective rays with vertex in the projection center. Finally $\bar{\Omega}$ is mapped into $\hat{\Omega}$ by collineation \mathbf{H}_c . The conic curve $\hat{\Omega}$ is the catadioptric image of the original line.

There are two kinds of degenerate conics: a locus of points consisting of a pair of distinct lines, and a locus consisting in a single line. In the former the rank of the corresponding 3×3 matrix is 2 and in the last the rank is 1. Since \mathbf{H}_c is a projective transformation then it is a full rank matrix and the line image $\hat{\Omega}$ is degenerate if, and only if, the corresponding conic $\bar{\Omega}$ is also degenerate (equation 3.2). Using the determinant of the matrix provided in equation 3.1 we conclude that $\bar{\Omega}$ is rank deficient whenever $\xi = 0$ or $n_z = 0$. If $\xi = 0$ then the sensor is a planar catadioptric system (see Tab. 2.3). The mapping becomes similar to conventional perspective cameras where a line in the scene is projected into a line. If $n_z = 0$ then the imaged line is coplanar with the Z-axis of the catadioptric reference frame, and the corresponding catadioptric image is also a line. This can be easily understood using the sphere model for the mapping (see Fig. 3.1).

3.2 The Conic Curve $\bar{\Omega}$

From equation 3.2 results that if $\mathbf{H}_c = \mathbf{I}$ then the catadioptric image of a line is $\hat{\Omega} = \bar{\Omega}$. The present section ignores the collineation \mathbf{H}_c and focuses on the effects of the non linear mapping \bar{h} in the catadioptric line projection. Both affine and euclidean geometry of the conic curve $\bar{\Omega}$ are studied.

3.2.1 Affine Geometry of $\bar{\Omega}$

Assuming $\mathbf{H}_c = \mathbf{I}$ then any line contained in plane $\Pi = (\mathbf{n}^t, 0)^t$, going through the projection center of the catadioptric system, is imaged into the conic $\bar{\Omega}$ (see Fig. 3.2). The affine characterization of $\bar{\Omega}$ is performed by assuming that the

3. Central Catadioptric Line Projection

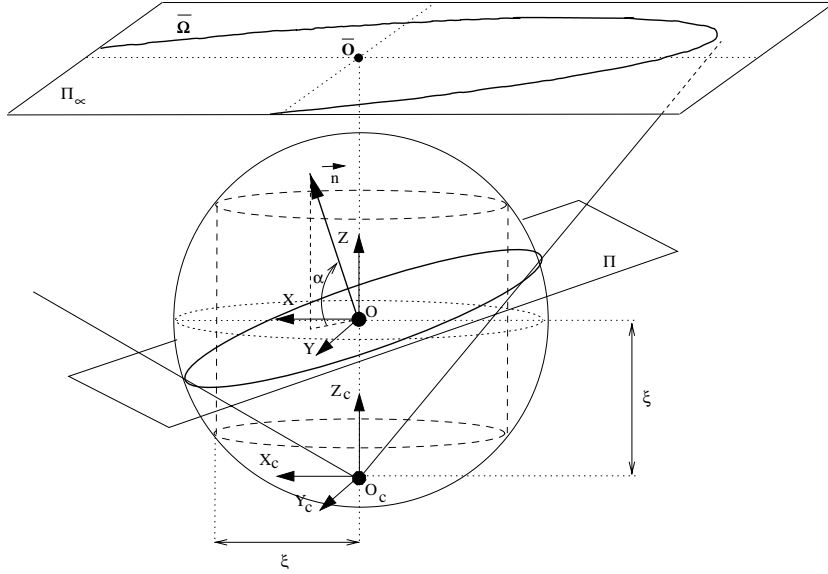


Figure 3.2: The central catadioptric image of a line when $\mathbf{H}_c = \mathbf{I}$

line at infinity $\bar{\pi}_\infty$ is in the canonical position ($\bar{\pi}_\infty = (0, 0, 1)^t$) [29, 39, 30]. The center is the pole of $\bar{\pi}_\infty$ with respect to the conic [62, 59]. The center \bar{C} of conic $\bar{\Omega}$ is computed by $\bar{C} = \bar{\Omega}^* \bar{\pi}_\infty$, with $\bar{\Omega}^*$ the conic envelope of $\bar{\Omega}$ (see Tab. 3.1).

The conic curve cuts the line at infinity in two points \bar{Q} and \bar{R} provided in Tab. 3.1. If the points are real and distinct, coincident or conjugate complex, then the conic $\bar{\Omega}$ is an hyperbola, a parabola, or an ellipse/circle [62, 59]. Notice that \bar{Q} and \bar{R} are respectively real and distinct, coincident or complex conjugate, if, and only if, the polynomial Δ of equation 3.3 is greater than, equal to, or less than zero. Δ is called the conic discriminant. For $\Delta > 0$, $\Delta = 0$ and $\Delta < 0$ the conic curve $\bar{\Omega}$ is respectively an hyperbola, a parabola, or an ellipse/circle.

$$\Delta = (n_x^2 + n_y^2)(1 - \xi^2) - n_z^2 \xi^2 \quad (3.3)$$

Consider the normal $\mathbf{n} = (n_x, n_y, n_z)^t$ to plane Π (see Fig. 3.2) and the angle α between \mathbf{n} and plane XOY (equation 3.4). From equation 3.3 and 3.4 it arises that $\Delta = 0$ whenever $\tan(\alpha)^2 = (1 - \xi^2)/\xi^2$. If the normal to the plane Π intersects the unit sphere in the dashed circles of Fig. 3.2 then the line image is a parabola. Moreover if the intersection point is between the circles the line image is an hyperbola and if the intersection point is above or below the circles the line

3.3. Projective Properties of the Central Catadioptric Line Image

Center	$\bar{C} = (n_x n_z, n_y n_z, -\Delta)^t$
Intersection with $\bar{\pi}_\infty$	$\bar{Q} = (1, \frac{n_x n_y (1-\xi^2) + n_z \xi^2 \sqrt{\Delta}}{n_z^2 \xi^2 - n_y^2 (1-\xi^2)}, 0)^t$; $\bar{R} = (1, \frac{n_x n_y (1-\xi^2) - n_z \xi^2 \sqrt{\Delta}}{n_z^2 \xi^2 - n_y^2 (1-\xi^2)}, 0)^t$

Table 3.1: Affine geometry of the conic curve $\bar{\Omega}$.

image is an ellipse/circle.

$$\alpha = \arctan\left(\frac{n_z}{\sqrt{n_x^2 + n_y^2}}\right) \quad (3.4)$$

3.2.2 Euclidean Geometry of $\bar{\Omega}$

For the affine characterization we have constrained the position of the line at infinity. We now restrict the representation further by assuming that the circular points \bar{I}_∞ and \bar{J}_∞ are in the canonical position ($\bar{I}_\infty = (1, i, 0)^t$ and $\bar{J}_\infty = (1, -i, 0)^t$). Consider the degenerate line conic $\bar{\omega}_\infty^*$ which consists of the two circular points ($\bar{\omega}_\infty^* = \bar{I}_\infty \bar{J}_\infty^t + \bar{J}_\infty \bar{I}_\infty^t$). Two lines in the euclidean plane are orthogonal if, and only if, they are conjugate with respect to $\bar{\omega}_\infty^*$ [39, 59, 30, 29].

A diameter of $\bar{\Omega}$ is a line going through the center of the conic. The principal axes are a pair of orthogonal diameters which are conjugate with respect to $\bar{\Omega}$. In general a conic only has a pair of principal axes (the exception is the circle). Any line containing one of the circular points \bar{I}_∞ or \bar{J}_∞ is called isotropic. A point \bar{F} is a focus of $\bar{\Omega}$ if, and only if, both isotropic lines $\bar{F}\bar{I}_\infty$ and $\bar{F}\bar{J}_\infty$ are tangent to $\bar{\Omega}$. In general a conic has four foci, two of which are real and two conjugate complex [62, 59]. Tab. 3.2 summarizes the euclidean parameters of the catadioptric line image $\bar{\Omega}$.

3.3 Projective Properties of the Central Catadioptric Line Image

In the central projection model derived in the previous chapter, points in the scene are projected onto the surface of an unit sphere centered in the effective viewpoint

3. Central Catadioptric Line Projection

Principal Axes	$\bar{\boldsymbol{\mu}} = (-n_y, n_x, 0)^t$; $\bar{\boldsymbol{\nu}} = (n_x\Delta, n_y\Delta, (n_x^2 + n_y^2)n_z)^t$
Major and Minor Axes	$\frac{\sqrt{n_x^2+n_y^2+n_z^2}}{ \Delta }$; $\frac{n_z^2\sqrt{n_x^2+n_y^2+n_z^2}\xi^2}{\Delta^2}$
Foci	$\bar{\boldsymbol{F}}_1 = \left(\frac{n_x}{n_z + \sqrt{(n_x^2+n_y^2+n_z^2)(1-\xi^2)}}, \frac{n_y}{n_z + \sqrt{(n_x^2+n_y^2+n_z^2)(1-\xi^2)}}, 1 \right)^t$; $\bar{\boldsymbol{F}}_2 = \left(\frac{n_x}{n_z - \sqrt{(n_x^2+n_y^2+n_z^2)(1-\xi^2)}}, \frac{n_y}{n_z - \sqrt{(n_x^2+n_y^2+n_z^2)(1-\xi^2)}}, 1 \right)^t$; $\bar{\boldsymbol{F}}_3 = \left(\frac{n_x n_z - i n_y \sqrt{(n_x^2+n_y^2+n_z^2)(1-\xi^2)}}{\Delta}, \frac{n_y n_z + i n_x \sqrt{(n_x^2+n_y^2+n_z^2)(1-\xi^2)}}{\Delta}, 1 \right)^t$; $\bar{\boldsymbol{F}}_4 = \left(\frac{n_x n_z + i n_y \sqrt{(n_x^2+n_y^2+n_z^2)(1-\xi^2)}}{\Delta}, \frac{n_y n_z - i n_x \sqrt{(n_x^2+n_y^2+n_z^2)(1-\xi^2)}}{\Delta}, 1 \right)^t$

Table 3.2: Euclidean geometry of the conic curve $\bar{\Omega}$

O. The catadioptric image is captured by a perspective camera which projects the points from the sphere onto a plane. If $\mathbf{H}_c = \mathbf{I}$ then the image plane is on the canonical position (by canonical we mean orthogonal to the forward looking Z axis). Since we are considering a projective framework there is invariance to scale changes. As a result it can be assumed that, whenever $\mathbf{H}_c = \mathbf{I}$, the image plane is the plane at infinity Π_∞ . To each sphere point \mathbf{P} corresponds a projective ray going through the camera center \mathbf{O}_c . Point \mathbf{P} projects on $\bar{\mathbf{P}}$ which is the intersection of the projective ray $\mathbf{O}_c\mathbf{P}$ with Π_∞ .

Consider a line in the scene which lies in plane Π as depicted in Fig.3.3. Π , going through the effective viewpoint \mathbf{O} with normal \mathbf{n} , intersects the spherical surface in a great circle. Points on the great circle define a central cone of projective rays with vertex in \mathbf{O}_c . The central cone of projective rays intersects Π_∞ in the conic curve $\bar{\Omega}$ (equation 3.1). Moreover a pencil of parallel planes intersects Π_∞ in the same line (the horizon line) and a pencil of parallel lines intersects Π_∞ in the same point (the direction point). Returning to Fig. 3.3, $\bar{\pi}$ is the horizon line of plane Π , $\bar{\mathbf{D}}$ is the direction point of line $\mathbf{P}_1\mathbf{P}_2$, and $\bar{\mathbf{N}}$ is the direction orthogonal to Π . Notice that space line $\mathbf{P}_1\mathbf{P}_2$ lies on plane Π thus $\bar{\pi}^t\bar{\mathbf{D}} = 0$. Moreover lines $\mathbf{P}_1\mathbf{P}_2$, $\mathbf{O}_c\mathbf{O}$, projective rays $\mathbf{O}_c\mathbf{P}_1$, $\mathbf{O}_c\mathbf{P}_2$ and the normal \mathbf{n} are coplanar, thus the corresponding direction points $\bar{\mathbf{D}}$, $\bar{\mathbf{O}}$, $\bar{\mathbf{P}}_1$, $\bar{\mathbf{P}}_2$ and $\bar{\mathbf{N}}$ are all collinear.

In section 3.2 we have assumed that $\mathbf{H}_c = \mathbf{I}$. Collineation \mathbf{H}_c depends on

3.3. Projective Properties of the Central Catadioptric Line Image

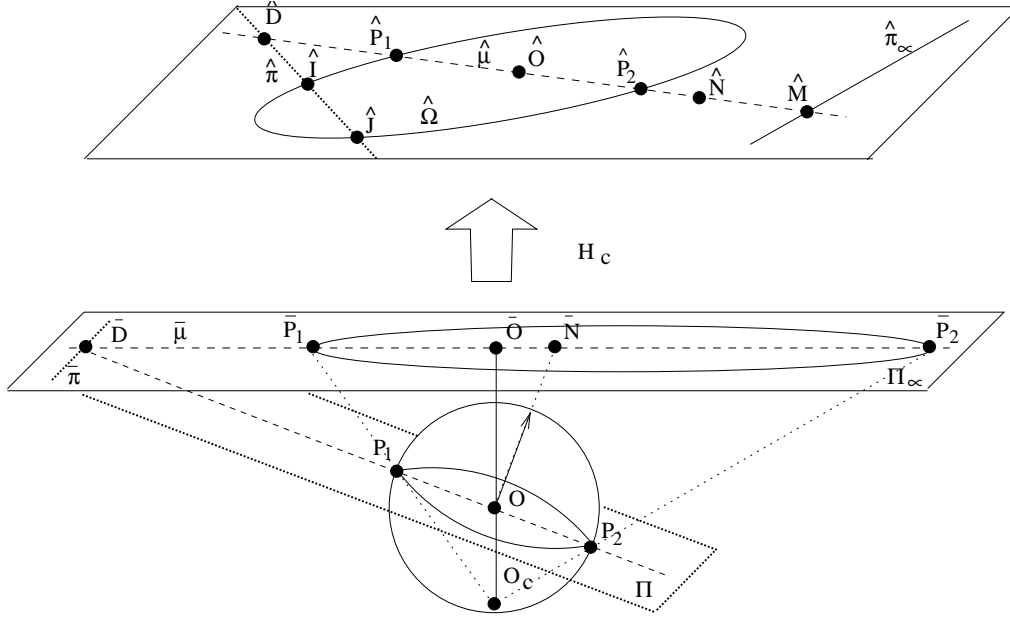


Figure 3.3: The central catadioptric image of a single line

camera intrinsic parameters, the relative rotation between the imaging device and the reflective surface, and mirror parameters (equation 2.7). In general $\mathbf{H}_c \neq \mathbf{I}$ and the final catadioptric image and Π_∞ are related by a general projective transformation between planes (Fig. 3.3). A generic point $\bar{\mathbf{P}}$ is mapped in $\hat{\mathbf{P}} = \mathbf{H}_c \bar{\mathbf{P}}$, the conic curve $\bar{\Omega}$ is imaged in $\hat{\Omega} = \mathbf{H}_c^{-t} \bar{\Omega} \mathbf{H}_c^{-1}$, and the line $\bar{\pi}$ is transformed in $\hat{\pi} = \mathbf{H}_c^{-t} \bar{\pi}$ [39, 29].

The projective transformation \mathbf{H}_c can change both the position of the line at infinity $\bar{\pi}_\infty$ and the circular points $\bar{\mathbf{I}}_\infty$ and $\bar{\mathbf{J}}_\infty$ [39, 30, 59]. In general the affine and euclidean parameters derived in section 3.2 do not hold. The circular points are fixed if, and only if, \mathbf{H}_c is a similarity transformation. If this is not the case, lines $\hat{\mu}$, $\hat{\nu}$ and points $\hat{\mathbf{F}}_1$, $\hat{\mathbf{F}}_2$, $\hat{\mathbf{F}}_3$, $\hat{\mathbf{F}}_4$ are no longer the principal axes and foci of the conic curve $\hat{\Omega}$ (see Tab. 3.2). Moreover point $\hat{\mathbf{C}} = \mathbf{H}_c \bar{\mathbf{C}}$ is the center of $\hat{\Omega}$ and the conic type is preserved if, and only if, \mathbf{H}_c is an affine transformation. Nevertheless the projective transformation preserves collinearity, incidence and the cross-ratio. These invariants are used to derive properties of central catadioptric line image that hold whenever $\mathbf{H}_c \neq \mathbf{I}$. The properties apply to any line projected on a non-degenerate conic. Section 3.3.1 establishes properties for single line imaging.

3. Central Catadioptric Line Projection

The propositions enunciated on section 3.3.2 are only valid for central parabolic systems ($\xi = 1$). Section 3.3.3 studies the central catadioptric projection of a pair of lines.

3.3.1 Catadioptric Projection of a Single Line

Fig. 3.3 is a scheme of the central catadioptric projection of a line lying on a plane $\Pi = (n_x, n_y, n_z, 0)^t$ going through the effective viewpoint \mathbf{O} . The plane intersects the sphere on a great circle which is projected on conic $\bar{\Omega}$ at the infinity plane Π_∞ . The point conic $\bar{\Omega}$ is transformed by \mathbf{H}_c in $\hat{\Omega}$ which is the the central catadioptric image of the original line (equation 3.2). The principal point $\bar{\mathbf{O}} = (0, 0, 1)^t$ and the normal direction $\bar{\mathbf{N}} = (n_x, n_y, n_z)^t$ of plane Π are mapped on points $\hat{\mathbf{O}}$ and $\hat{\mathbf{N}}$ by collineation \mathbf{H}_c . Plane Π intersects Π_∞ on the horizon line $\bar{\pi}$ which is mapped on $\hat{\pi}$ at the catadioptric image plane. Moreover we will consider the absolute conic $\bar{\Omega}_\infty$ (not depicted) which is transformed in the conic $\hat{\Omega}_\infty$ under the projectivity \mathbf{H}_c .

Proposition 3.1: *If the point conic $\hat{\Omega}$ is a line image then the polar of the image center $\hat{\mathbf{O}}$ with respect to $\hat{\Omega}$ is the horizon line $\hat{\pi}$ of the plane Π containing the imaged line and the effective viewpoint ($\hat{\pi} = \hat{\Omega}.\hat{\mathbf{O}}$).*

Proof: The plane Π , containing the imaged line, intersects the unit sphere in a great circle. Consider a generic line, lying on Π and going through the effective viewpoint \mathbf{O} , with direction point $\bar{\mathbf{D}}$ (Fig. 3.3). The line intersects the great circle in two antipodal points \mathbf{P}_1 and \mathbf{P}_2 which are equidistant to \mathbf{O} . This implies that, in the plane at infinity, points $\bar{\mathbf{D}}$ and $\bar{\mathbf{O}}$ are harmonic with respect to $\bar{\mathbf{P}}_1$ and $\bar{\mathbf{P}}_2$. Points $\bar{\mathbf{P}}_1$ and $\bar{\mathbf{P}}_2$ lie in the conic curve $\bar{\Omega}$, and the locus of direction points $\bar{\mathbf{D}}$ is the horizon line $\bar{\pi}$ of plane Π . Thus $\bar{\pi}$ is the polar of $\bar{\mathbf{O}}$ with respect to $\bar{\Omega}$. The catadioptric image is related with the plane at infinity by a collineation \mathbf{H}_c . The proposition is proved since pole/polar relations are invariant under a projective transformation.

Proposition 3.2: *The absolute conic $\bar{\Omega}_\infty$ is mapped on the point conic $\hat{\Omega}_\infty$ at the catadioptric image plane. If $\hat{\Omega}$ is the catadioptric image of a line, then the polar*

3.3. Projective Properties of the Central Catadioptric Line Image

line $\hat{\pi}$ of the image center $\hat{\mathbf{O}}$ ($\hat{\pi} = \hat{\mathbf{O}} \cdot \hat{\mathbf{O}}$) intersects the conic locus $\hat{\mathbf{Q}}$ in two points $\hat{\mathbf{I}}$ and $\hat{\mathbf{J}}$ which lie on $\hat{\mathbf{Q}}_\infty$.

Proof: In Fig. 3.3 the plane Π intersects the unit sphere on a great circle. The great circle defines a central cone of projective rays with vertex in \mathbf{O}_c . The central cone is a quadric surface \mathbf{S} in space. In general a quadric intersects a plane in a conic curve. The quadric surface \mathbf{S} intersects the plane at infinity Π_∞ in the conic curve $\bar{\mathbf{Q}}$. The conic curve $\bar{\mathbf{Q}}$ and the absolute conic $\bar{\mathbf{Q}}_\infty$, both lying in Π_∞ , intersect in four points. Each pair of intersection points $\bar{\mathbf{I}}$ and $\bar{\mathbf{J}}$ defines a line which is the horizon line of a pencil of parallel planes (real or complex). These planes intersect the original quadric \mathbf{S} in conic sections. Notice that both points $\bar{\mathbf{I}}$ and $\bar{\mathbf{J}}$ lie in each one of these intersection conics. They lie simultaneously in the quadric surface \mathbf{S} (they belong to $\bar{\mathbf{Q}}$) and in the cutting plane (they belong to the corresponding horizon line). Points $\bar{\mathbf{I}}$ and $\bar{\mathbf{J}}$ also belong to the absolute conic $\bar{\mathbf{Q}}_\infty$ and are circular points. Thus, one concludes that each plane of the defined pencil intersects the central cone \mathbf{S} on a circular section (for further details see [63, 62]). Moreover if a plane Π intersects the quadric surface \mathbf{S} in a circular section then its horizon line must go through two intersection points of $\bar{\mathbf{Q}}$ with $\bar{\mathbf{Q}}_\infty$. The established relations hold in the catadioptric image plane after the projective transformation \mathbf{H}_c .

Proposition 3.3: Consider the line $\hat{\mu}$ defined by the principal point $\hat{\mathbf{O}}$ and the normal direction $\hat{\mathbf{N}}$ in the catadioptric image plane ($\hat{\mu} = \hat{\mathbf{O}} \wedge \hat{\mathbf{N}}$). Line $\hat{\mu}$ is the locus where the major axis $\bar{\mu}$ of the conic curve $\bar{\mathbf{Q}}$ is mapped by collineation \mathbf{H}_c .

Proof: Accordingly to Tab. 3.2 the major axis of the conic curve $\bar{\mathbf{Q}}$ is $\bar{\mu} = (-n_y, n_x, 0)^t$. Both the principal point $\bar{\mathbf{O}} = (0, 0, 1)^t$ and the normal direction $\bar{\mathbf{N}} = (n_x, n_y, n_z)^t$ lie on $\bar{\mu}$ ($\bar{\mu}^t \bar{\mathbf{O}} = \bar{\mu}^t \bar{\mathbf{N}} = 0$). Fig. 3.3 shows that the major axis $\bar{\mu}$ is the intersection line between Π_∞ and the plane containing both the normal direction \mathbf{n} and the Z axis. Since collineation \mathbf{H}_c preserves incidence and collinearity then both points $\hat{\mathbf{O}}$ and $\hat{\mathbf{N}}$ must lie on the locus $\hat{\mu}$ where $\bar{\mu}$ is mapped. However remember that in general line $\hat{\mu}$ is no longer the major axis of the catadioptric line image $\hat{\mathbf{Q}}$.

Consider line $\bar{\pi}_\infty = (0, 0, 1)^t$ lying on plane Π_∞ . Line $\bar{\pi}_\infty$ is mapped on line $\hat{\pi}_\infty$

3. Central Catadioptric Line Projection

by projective transformation \mathbf{H}_c . Line $\hat{\pi}_\infty$ is in the canonical position if, and only if, \mathbf{H}_c is an affine transformation. Notice that $\hat{\pi}_\infty$ is the intersection line between Π_∞ and the catadioptric image plane. Consider the intersection points $\bar{\mathbf{D}}$ of $\bar{\pi}$ with $\bar{\mu}$, and $\bar{\mathbf{M}}$ of $\bar{\pi}_\infty$ with $\bar{\mu}$. These points are mapped on points $\hat{\mathbf{D}}$ and $\hat{\mathbf{M}}$, which still are intersection points since \mathbf{H}_c preserves incidence (see Fig. 3.3).

Corollary 3.1: *The pole of line $\hat{\mu}$ with respect to the point conic $\hat{\Omega}$ lies on $\hat{\pi}_\infty$ which is the intersection line of the catadioptric image with Π_∞ .*

Proof: The major axis $\bar{\mu}$ is a diameter of the point conic $\bar{\Omega}$, thus the corresponding pole lies at the line at infinity $\bar{\pi}_\infty$ [59]. Since pole/polar relations are preserved under projective transformations it comes that the pole of $\hat{\mu}$ (point $\hat{\Omega}^* \cdot \hat{\mu}$ with $\hat{\Omega}^*$ the conic envelope of $\hat{\Omega}$) lies on $\hat{\pi}_\infty$.

Proposition 3.4: *The cross ratio between points $\hat{\mathbf{O}}$, $\hat{\mathbf{N}}$, $\hat{\mathbf{D}}$ and $\hat{\mathbf{M}}$, lying on $\hat{\mu}$, only depends on the angle between plane Π and plane XOY of the catadioptric reference frame. In particular $\{\hat{\mathbf{M}}, \hat{\mathbf{O}}; \hat{\mathbf{N}}, \hat{\mathbf{D}}\} = -\tan(\alpha)^2$ with α the angle of equation 3.4.*

Proof: We have already seen that the principal point in Π_∞ is $\bar{\mathbf{O}} = (0, 0, 1)^t$ and the normal direction is $\bar{\mathbf{N}} = (n_x, n_y, n_z)^t$. Moreover $\bar{\pi}_\infty = (0, 0, 1)^t$, $\bar{\mu} = (-n_y, n_x, 0)^t$ and the conic $\bar{\Omega}$ is given by equation 3.1. From proposition 1 comes that the horizon line of plane Π is $\bar{\pi} = \bar{\Omega}\bar{\mathbf{O}}$. The intersection point of lines $\bar{\pi}$ and $\bar{\mu}$ is $\bar{\mathbf{D}} = \bar{\mu} \wedge \bar{\pi} = (-n_x n_z, -n_y n_z, n_x^2 + n_y^2)^t$. Point $\bar{\mathbf{M}}$ is the direction point of $\bar{\mu}$, thus $\bar{\mathbf{M}} = \bar{\mu} \wedge \bar{\pi}_\infty = (n_x, n_y, 0)^t$. Computing the cross-ratio between the four points arises $\{\bar{\mathbf{M}}, \bar{\mathbf{O}}; \bar{\mathbf{N}}, \bar{\mathbf{D}}\} = -\tan(\alpha)^2$ with α given by equation 3.4. The cross-ratio is a projective invariant and the proposition is proved.

Point $\bar{\mathbf{C}}$ is the center of the conic $\bar{\Omega}$ lying on plane Π_∞ (see Tab. 3.1). By definition $\bar{\mathbf{C}}$ is the pole of the line at infinity $\bar{\pi}_\infty$ with respect to $\bar{\Omega}$ [59, 62]. The pole/polar relations are preserved by projective transformations. Thus, if $\bar{\mathbf{C}}$ is mapped on points $\hat{\mathbf{C}}$ in the catadioptric image then it comes that $\hat{\mathbf{C}} = \hat{\Omega}^* \cdot \hat{\pi}_\infty$. Moreover $\hat{\mathbf{C}}$ must lie on line $\hat{\mu}$ since the major axis $\bar{\mu}$ goes through the center $\bar{\mathbf{C}}$.

3.3. Projective Properties of the Central Catadioptric Line Image

Proposition 3.5: *The cross ratio between points \hat{O} , \hat{N} , \hat{D} and \hat{C} , lying on $\hat{\mu}$, depends only on the shape of the reflective surface used in the central catadioptric system. In particular $\{\hat{O}, \hat{D}; \hat{N}, \hat{C}\} = \xi^2$.*

Proof: The coordinates of points \bar{O} , \bar{D} and \bar{N} have already been derived and center \bar{C} is provided on Tab. 3.1. Computing the cross-ratio comes that $\{\bar{O}, \bar{D}; \bar{N}, \bar{C}\} = \xi^2$. The proposition is proved since the collineation H_c preserves the cross-ratio.

Corollary 3.2: *The polar lines of points \hat{O} and \hat{N} , with respect to the image of the absolute conic $\hat{\Omega}_\infty$, are $\hat{\pi}_\infty$ and $\hat{\pi}$ ($\hat{\pi}_\infty = \hat{\Omega}_\infty \hat{O}$ and $\hat{\pi} = \hat{\Omega}_\infty \hat{N}$).*

Proof: \bar{N} is the normal direction of plane Π , which intersects Π_∞ on line $\bar{\pi}$. Thus, the polar of \bar{N} with respect to the absolute conic $\bar{\Omega}_\infty$ is the horizon line $\bar{\pi}$. The property is preserved by projective transformation H_c . In a similar manner the pole/polar relation between \hat{O} and $\hat{\pi}_\infty$ is proved taking into account that plane Π_∞ is orthogonal to the Z-axis of the catadioptric reference frame (see Fig. 3.3).

Consider \hat{N}^* lying on line $\hat{\mu}$ in the catadioptric image plane. Points \hat{N} and \hat{N}^* are conjugate with respect to the catadioptric line image $\hat{\Omega}$ ($\hat{N}^t \hat{\Omega} \hat{N}^* = 0$). This point is the locus where $\bar{N}^* = (\bar{\Omega} \bar{N}) \wedge \bar{\mu}$ is mapped by collineation H_c . Equation 3.5 is a relation of cross-ratios between points \hat{O} , \hat{M} , \hat{N} , \hat{C} and \hat{N}^* . The result can be proved in a similar manner as proposition 3.4 and 3.5, taking into account that cross-ratios are projective invariants and that $\bar{N}^* = (n_x n_z, n_y n_z, -(1 - \xi^2)(n_x^2 + n_y^2))^t$. The established relationship will be useful later for calibration purposes.

$$\{\hat{C}, \hat{N}; \hat{M}, \hat{O}\} = \xi^2 - \frac{2\xi^2 \{\hat{N}^*, \hat{N}; \hat{M}, \hat{C}\}}{(1 - \xi^2) \left(1 + \sqrt{1 + \frac{4\xi^2 \{\hat{N}^*, \hat{N}; \hat{M}, \hat{C}\}}{(1 - \xi^2)^2}}\right)} \quad (3.5)$$

3.3.2 Paracatadioptric Projection of a Single Line

The propositions enunciated in this subsection are only valid for central catadioptric systems combining a parabolic mirror with an orthographic camera (Fig. 2.1). If the mirror is parabolic then the ξ parameter is unitary (Tab. 2.3). Replacing ξ

3. Central Catadioptric Line Projection

by 1 on equation 3.1 yields

$$\bar{\Omega} = \begin{bmatrix} -n_z^2 & 0 & n_x n_z \\ 0 & -n_z^2 & n_y n_z \\ n_x n_z & n_y n_z & n_z^2 \end{bmatrix} \quad (3.6)$$

Consider the scheme of Fig. 3.3 for the central catadioptric projection of a single line. Plane Π , containing both the line and the effective viewpoint O , intersects the sphere in a great circle. For the particular situation of the paracatadioptric camera the parameter $\xi = 1$ and the re-projection center O_c lies on the sphere surface. The mapping from the sphere to the plane at infinity Π_∞ is a stereographic projection [34]. The stereographic projection maps any circle in the sphere into a circle in the plane [59]. Thus the great circle is projected into a circle $\bar{\Omega}$ lying on plane Π_∞ (equation 3.6). The paracatadioptric image of the line is $\hat{\Omega} = H_c^{-t} \bar{\Omega} H_c^{-1}$, with $H_c = K_c R_c M_c$ (equations 3.2 and 2.7). Since the system is parabolic then the optical axis of the orthographic camera must be aligned with the symmetry axis of the mirror and $R_c = I$ (Tab. 2.2). The transformation H_c is always affine. Since an affine transformation does not change the type of conic, then the paracatadioptric line image $\hat{\Omega}$ is always a circle/ellipse.

Consider the following points lying on plane Π_∞ : $\bar{I}_\infty = (1, i, 0)^t$, $\bar{J}_\infty = (1, -i, 0)^t$, $\bar{G} = (1, 0, -i)^t$ and $\bar{H} = (1, 0, i)^t$. Points \bar{I}_∞ and \bar{J}_∞ are the circular points of the plane. It is well known that any circle must go through the circular points. Since $\bar{\Omega}$ is always a circle, then it is true that $\bar{I}_\infty^t \bar{\Omega} \bar{I}_\infty = 0$ and $\bar{J}_\infty^t \bar{\Omega} \bar{J}_\infty = 0$. Moreover from equation 3.6 arises that $\bar{G}^t \bar{\Omega} \bar{H} = 0$. Thus points \bar{G} and \bar{H} are conjugate with respect to the conic curve $\bar{\Omega}$. Assume that collineation H_c maps points \bar{I}_∞ , \bar{J}_∞ , \bar{G} and \bar{H} into points \hat{I}_∞ , \hat{J}_∞ , \hat{G} and \hat{H} in the paracatadioptric image plane.

Proposition 3.6: *A conic curve $\hat{\Omega}$ is the paracatadioptric image of a line in the scene if, and only if, it contains points \hat{I}_∞ and \hat{J}_∞ , and points \hat{G} , \hat{H} are harmonic conjugate with respect to $\hat{\Omega}$.*

Proof: The circle $\bar{\Omega}$ must go through the plane circular points \bar{I}_∞ , \bar{J}_∞ [59, 39]. Since collineation H_c preserves incidence, then both points \hat{I}_∞ , \hat{J}_∞ lie on the parabolic line image $\hat{\Omega}$. Moreover the projective transformation also preserves

3.3. Projective Properties of the Central Catadioptric Line Image

the cross-ratio and pole/polar relations. Since points \bar{G} , \bar{H} are conjugate with respect to $\bar{\Omega}$, then \hat{G} , \hat{H} are also conjugate with respect to $\hat{\Omega}$. Thus, if $\hat{\Omega}$ is a paracatadioptric line image, then it must verify $\hat{I}_\infty^t \hat{\Omega} \hat{I}_\infty = 0$, $\hat{J}_\infty^t \hat{\Omega} \hat{J}_\infty = 0$ and $\hat{G}^t \hat{\Omega} \hat{H} = 0$. The derived conditions are necessary, nevertheless it is not clear that they are sufficient. By sufficient we mean that if a conic curve in the paracatadioptric image plane verifies these 3 constraints, then it is the locus where a certain line in the scene is projected. Notice that, neglecting the scale factor, the conic curve $\bar{\Omega}$ provided by equation 3.6 is a function of 2 independent parameters. These 2 degrees of freedom (DOF) are associated with the pose of plane Π containing the line and the effective viewpoint (Fig. 3.3). Since in general a conic curve has 5 DOF, then we must be able to find 3, and no more than 3, independent constraints. This proves the sufficiency of the statement.

Corollary 3.3: *In a central parabolic vision system all line images intersect in two points \hat{I}_∞ and \hat{J}_∞ lying on $\hat{\pi}_\infty$.*

Proof: The proof of this corollary is straightforward. The circular points of plane Π_∞ lie on line $\bar{\pi}_\infty$. Since the projective transformation preserves incidence and collinearity, then points \hat{I}_∞ , \hat{J}_∞ lie on line $\hat{\pi}_\infty$ which is the locus where $\bar{\pi}_\infty$ is mapped. From proposition 3.6 arises that all paracatadioptric images $\hat{\Omega}$ must go through the image of the circular points which completes the proof.

Proposition 3.7: *In a central parabolic image, point \hat{N} and line $\hat{\pi}_\infty$ are pole/polar with respect to the corresponding line image $\hat{\Omega}$.*

Proof: Plane Π , with normal direction \bar{N} , intersects the unit sphere in a great circle which generates the conic $\bar{\Omega}$ (Fig. 3.3). Point \bar{C} , provided in Tab. 3.1, is the center of the conic curve. Making $\xi = 1$ comes that $\bar{C} = \bar{N}$. For a parabolic system the center of conic $\bar{\Omega}$ is the normal direction of the plane Π containing the imaged line. Since the conic center is the pole of the line at infinity, then point \bar{N} and line π_∞ are pole/polar with respect to $\bar{\Omega}$. Pole/polar relations are projective invariants and the statement is proved.

3. Central Catadioptric Line Projection

3.3.3 Catadioptric Projection of a Pair of Lines

Fig. 3.4 depicts the central catadioptric projection of a pair of lines. Two lines, lying on planes Π_i and Π_j , are imaged on conics $\hat{\Omega}_i$ and $\hat{\Omega}_j$. The schematic is similar to the one used on Fig. 3.3 for a single line. Configurations leading to degenerate conics are excluded. Thus the imaged lines are not coplanar with the z-axis and the central catadioptric system is not a perspective camera.

The planes Π_i and Π_j , going through the effective viewpoint O , cut the unit sphere in two great circles which intersect each other in two antipodal points (F_{ij} and B_{ij}). The direction point of line $F_{ij}B_{ij}$, going through the antipodal points, is \bar{D}_{ij} . The plane Φ_{ij} , orthogonal to direction \bar{D}_{ij} and containing the origin O , intersects the infinity plane Π_∞ on line $\bar{\eta}_{ij}$. The direction \bar{D}_{ij} is common to planes Π_i and Π_j . Their normal vectors belong to plane Φ_{ij} and the corresponding direction points \bar{N}_i and \bar{N}_j , lie on the horizon line $\bar{\eta}_{ij}$.

Each great circle defines a central cone of projective rays with vertex in O_c , which intersects the plane at infinity Π_∞ in a conic curve. The conics associated with Π_i and Π_j are respectively $\bar{\Omega}_i$ and $\bar{\Omega}_j$. The conic curves intersect each other in two real points \bar{F}_{ij} and \bar{B}_{ij} . These points are the projection of the antipodal points F_{ij} and B_{ij} . The line going through the two intersection points is $\bar{\mu}_{ij} = \bar{F}_{ij} \wedge \bar{B}_{ij}$.

Conics $\bar{\Omega}_i$ and $\bar{\Omega}_j$ are imaged in $\hat{\Omega}_i$ and $\hat{\Omega}_j$. These are the central catadioptric images of the original lines. Points \bar{F}_{ij} , \bar{B}_{ij} , \bar{D}_{ij} , \bar{N}_i and \bar{N}_j are transformed on points \hat{F}_{ij} , \hat{B}_{ij} , \hat{D}_{ij} , \hat{N}_i and \hat{N}_j in the catadioptric image plane. $\hat{\mu}_{ij}$ and $\hat{\eta}_{ij}$ are the locus where lines $\bar{\mu}_{ij}$ and $\bar{\eta}_{ij}$ are mapped. The projective transformation H_c preserves all the incidence and collinearity relations.

Proposition 3.8: *If \hat{F}_{ij} and \hat{B}_{ij} are the intersection points of two catadioptric line images $\hat{\Omega}_i$ and $\hat{\Omega}_j$, then the image center \hat{O} is always collinear with \hat{F}_{ij} and \hat{B}_{ij} .*

Proof: Consider the plane Λ_{ij} , defined by the two antipodal points F_{ij} , B_{ij} and the camera center O_c (Fig. 3.4). Notice that Λ_{ij} always contains the effective viewpoint O . The projective rays O_cF_{ij} and O_cB_{ij} lie on the plane and $\bar{\mu}_{ij}$ is the horizon line of Λ_{ij} . Moreover the projective ray O_cO , intersecting Π_∞ in \bar{O} , also belongs to Λ_{ij} . The direction point \bar{O} lies on $\bar{\mu}_{ij}$ and is collinear with \bar{F}_{ij} and

3.3. Projective Properties of the Central Catadioptric Line Image

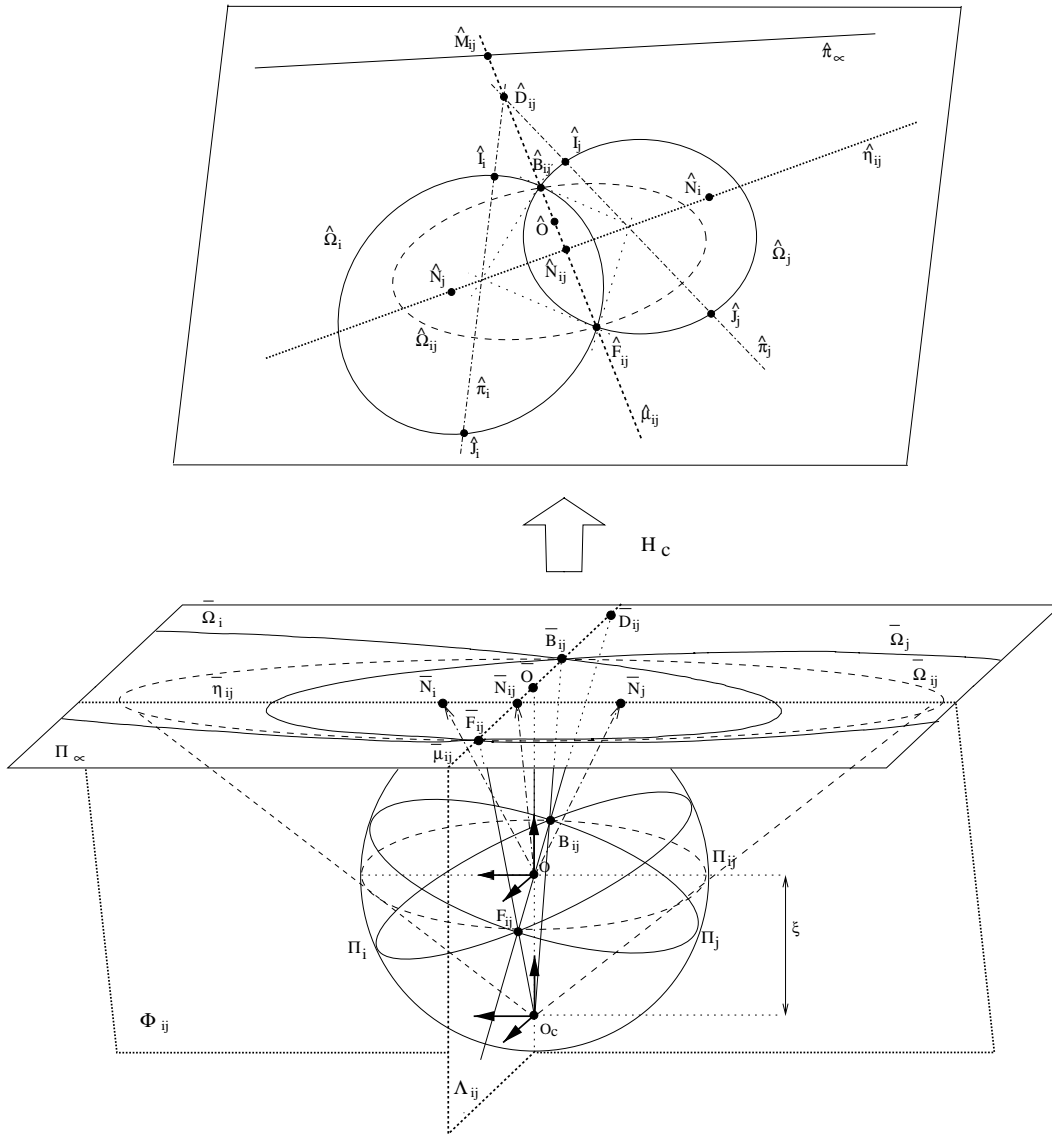


Figure 3.4: The central catadioptric image of a pair of lines

3. Central Catadioptric Line Projection

$\bar{\mathbf{B}}_{ij}$. The proposition is proved since the projective transformation \mathbf{H}_c preserves collinearity.

Corollary 3.4: *Consider two catadioptric line images $\hat{\Omega}_i$ and $\hat{\Omega}_j$, intersecting on points $\hat{\mathbf{F}}_{ij}$ and $\hat{\mathbf{B}}_{ij}$, and the image center $\hat{\mathbf{O}}$. If $\hat{\pi}_i, \hat{\pi}_j$ are the polar lines of $\hat{\mathbf{O}}$ with respect to $\hat{\Omega}_i, \hat{\Omega}_j$ and $\hat{\mu}_{ij}$ is the line going through the two intersection points then $\hat{\pi}_i, \hat{\pi}_j$ and $\hat{\mu}_{ij}$ intersect in the same point $\hat{\mathbf{D}}_{ij}$.*

Proof: We have already seen that direction $\bar{\mathbf{D}}_{ij}$ is common to planes Π_i, Π_j and Λ_{ij} (Fig. 3.4). Thus the corresponding vanishing lines $\bar{\pi}_i, \bar{\pi}_j$ and $\bar{\mu}_{ij}$ must intersect on $\bar{\mathbf{D}}_{ij}$. From proposition 3.1 comes that the polar lines $\hat{\pi}_i$ and $\hat{\pi}_j$ are the locus where the horizon lines of planes Π_i and Π_j are mapped. Moreover line $\bar{\mu}_{ij}$ in Π_∞ is transformed in $\hat{\mu}_{ij}$ at the catadioptric image plane. Since \mathbf{H}_c preserves incidence, comes that $\hat{\pi}_i, \hat{\pi}_j$ and $\hat{\mu}_{ij}$ must intersect in the same point $\hat{\mathbf{D}}_{ij}$, which is the locus where the common direction $\bar{\mathbf{D}}_{ij}$ is mapped.

Corollary 3.5: *Two parallel lines (non-coplanar with the center \mathbf{O}) are projected in the catadioptric image plane on conics $\hat{\Omega}_i$ and $\hat{\Omega}_j$. If the polar lines of $\hat{\mathbf{O}}$ are respectively $\hat{\pi}_i$ and $\hat{\pi}_j$, then the direction of the parallel lines is $\bar{\mathbf{D}}_{ij} = \mathbf{H}_c^{-1}(\hat{\pi}_i \wedge \hat{\pi}_j)$.*

Proof: If two lines are parallel then the common direction $\bar{\mathbf{D}}_{ij}$ of the corresponding planes Π_i and Π_j is the direction of the parallel lines (Fig. 3.4). The direction point $\bar{\mathbf{D}}_{ij}$ is mapped in $\hat{\mathbf{D}}_{ij}$ in the catadioptric image plane ($\hat{\mathbf{D}}_{ij} = \mathbf{H}_c \bar{\mathbf{D}}_{ij}$). This is the intersection point of polar lines $\hat{\pi}_i$ and $\hat{\pi}_j$, as stated in corollary 3.4 ($\hat{\mathbf{D}}_{ij} = \hat{\pi}_i \wedge \hat{\pi}_j$). From the above comes that the direction of the parallel lines is $\bar{\mathbf{D}}_{ij} = \mathbf{H}_c^{-1}(\hat{\pi}_i \wedge \hat{\pi}_j)$.

Planes Π_i and Π_j define two great circles in the unit sphere surface as depicted in Fig. 3.4. The great circles intersect each other on the antipodal points \mathbf{F}_{ij} and \mathbf{B}_{ij} . Plane Φ_{ij} is orthogonal to direction $\bar{\mathbf{D}}_{ij}$ of line $\mathbf{F}_{ij}\mathbf{B}_{ij}$, and intersects Π_∞ on line $\bar{\eta}_{ij}$. Consider the pencil of planes containing line $\mathbf{F}_{ij}\mathbf{B}_{ij}$. These planes go through the origin \mathbf{O} and the corresponding normal direction $\bar{\mathbf{N}}$ lies on $\bar{\eta}_{ij}$. Line $\bar{\eta}_{ij}$ is called the line of the normals of the pencil defined by Π_i and Π_j . Each plane

3.3. Projective Properties of the Central Catadioptric Line Image

of the pencil intersects the unit sphere in a great circle going through points F_{ij} and B_{ij} . Thus, any line lying on these planes is projected into a conic $\hat{\Omega}$ going through points \hat{F}_{ij} , \hat{B}_{ij} in the catadioptric image plane.

$\hat{\Omega}_i$ and $\hat{\Omega}_j$ are two point conics which are the catadioptric images of two lines lying on planes Π_i and Π_j (Fig. 3.4). Any pair of catadioptric line images has a line $\hat{\mu}_{ij}$ and $\hat{\eta}_{ij}$ associated with it. Consider the pencil of planes Π , defined by Π_i and Π_j , and the corresponding planes Λ_{ij} and Φ_{ij} . $\hat{\mu}_{ij}$ is the line going through the intersection points \hat{F}_{ij} , \hat{B}_{ij} of conics $\hat{\Omega}_i$ and $\hat{\Omega}_j$. The horizon line of plane Λ_{ij} is transformed in $\hat{\mu}_{ij}$ by collineation H_c . $\hat{\eta}_{ij}$ is the locus where the line of normals $\bar{\eta}_{ij}$ is mapped in the catadioptric image plane. The points \hat{N}_i and \hat{N}_j , associated with $\hat{\Omega}_i$ and $\hat{\Omega}_j$, lie in $\hat{\eta}_{ij}$. Moreover if a catadioptric line image $\hat{\Omega}$ goes through points \hat{F}_{ij} , \hat{B}_{ij} then the corresponding \hat{N} lies on $\hat{\eta}_{ij}$ and the imaged line is contained by a plane of the pencil defined by Π_i and Π_j . Notice that $\hat{\Omega}_{ij}$, depicted in Fig. 3.4, is a particular case of this family with the corresponding normal point \hat{N}_{ij} in the intersection of lines $\hat{\mu}_{ij}$ and $\hat{\eta}_{ij}$.

Proposition 3.9: *Consider the pair of catadioptric line images $\hat{\Omega}_i$ and $\hat{\Omega}_j$, intersecting on points \hat{F}_{ij} and \hat{B}_{ij} , and the corresponding lines $\hat{\mu}_{ij}$ and $\hat{\eta}_{ij}$. If conic $\hat{\Omega}$ is a line image going through points \hat{F}_{ij} , \hat{B}_{ij} then the pole of $\hat{\mu}_{ij}$ with respect to $\hat{\Omega}$ lies on $\hat{\eta}_{ij}$.*

Proof: The catadioptric line images $\hat{\Omega}$, $\hat{\Omega}_i$ and $\hat{\Omega}_j$ intersect in only two visible points. The line going through the intersection points \hat{F}_{ij} and \hat{B}_{ij} is $\hat{\mu}_{ij}$. Point \hat{Y} is the pole of $\hat{\mu}_{ij}$ with respect to $\hat{\Omega}$ ($\hat{Y} = \hat{\Omega}^* \cdot \hat{\mu}_{ij}$). The goal is to prove that point \hat{Y} lies on line $\hat{\eta}_{ij}$. The polar line of the image center \hat{O} with respect to $\hat{\Omega}$ is $\hat{\pi}$. Since \hat{O} lies on $\hat{\mu}_{ij}$ (proposition 3.8), then point \hat{Y} (the pole of $\hat{\mu}_{ij}$) must lie on $\hat{\pi}$ [59]. Line $\hat{\pi}$ intersects $\hat{\mu}_{ij}$ on point \hat{D}_{ij} , and the conic $\hat{\Omega}$ on points \hat{I} , \hat{J} . Since \hat{Y} and $\hat{\mu}_{ij}$ are pole/polar, then the pairs of points $\{\hat{I}, \hat{J}\}$ and $\{\hat{D}_{ij}, \hat{Y}\}$ are harmonic conjugates [59]. From proposition 3.1, comes that $\hat{\pi}$ is the locus where the horizon line of plane Π is mapped. Plane Π , containing the imaged line, belongs to the pencil defined by Π_i and Π_j (Fig. 3.4). The pencil common direction \bar{D}_{ij} is orthogonal to plane Φ_{ij} whose vanishing line is $\bar{\eta}_{ij}$. Thus, if the absolute conic is mapped on $\hat{\Omega}_\infty$ by H_c , then \hat{D}_{ij} and $\hat{\eta}_{ij}$ are pole/polar with respect to $\hat{\Omega}_\infty$. Line $\hat{\pi}$ intersects conics $\hat{\Omega}_\infty$ and $\hat{\Omega}$ on the same pair of points \hat{I} , \hat{J} (proposition 3.2).

3. Central Catadioptric Line Projection

Assume that lines $\hat{\pi}$ and $\hat{\eta}_{ij}$ intersect on point $\hat{\mathbf{K}}$. Since $\hat{\mathbf{D}}_{ij}$ and $\hat{\eta}_{ij}$ are pole/polar with respect to $\hat{\Omega}_\infty$, then the pairs of points $\{\hat{\mathbf{I}}, \hat{\mathbf{J}}\}$ and $\{\hat{\mathbf{D}}_{ij}, \hat{\mathbf{K}}\}$ are also harmonic conjugates. $\hat{\mathbf{Y}}$ and $\hat{\mathbf{K}}$ must be the same point and the proposition is proved.

3.4 Calibration with Minimal Information

The theory derived can be applied to the calibration of central catadioptric system using line images. Section 3.4.1 proves that, given the ξ parameter of the mirror and the relative pose of the camera and the reflective surface, it is possible to calibrate any hyperbolic/elliptical system from the image of two lines. Section 3.4.2 shows how to use three line images to calibrate any central panoramic system.

3.4.1 Calibrating an Hyperbolic/Elliptical System from Two Lines

The calibration strategy presented in this section is only valid for central catadioptric systems with an hyperbolic/elliptical mirror (Fig. 3.1). The ξ parameter, depending on mirror shape, is known. The plane at infinity Π_∞ intersects the catadioptric image plane on line $\hat{\pi}_\infty$ (Fig. 3.3 and 3.4). It is assumed that the position of $\hat{\pi}_\infty$ is known as well. On the majority of hyperbolic sensors commercially available the perspective camera is not rotated with relation to the reflective surface. The rotation matrix \mathbf{R}_c is the identity and collineation \mathbf{H}_c is an affine transformation (equation 2.7). Under this circumstances the line $\hat{\pi}_\infty$ is in the canonical position at the catadioptric image plane ($\hat{\pi}_\infty = (0, 0, 1)^t$).

Tab. 3.4 summarizes the algorithm to estimate the position of conic $\hat{\Omega}_\infty$ from the catadioptric image of two lines. The lines can not be coplanar with the mirror symmetry axis, otherwise they would be imaged in a degenerate conic (Fig. 3.1). This is the unique constraint on the line position in the scene. Conic $\hat{\Omega}_\infty$ is the locus where the absolute conic is mapped ($\hat{\Omega}_\infty = \mathbf{H}_c^{-t} \mathbf{H}_c^{-1}$). If the camera is not rotated with relation to the mirror ($\mathbf{R}_c = \mathbf{I}$), then collineation \mathbf{H}_c can be recovered from the Cholesky decomposition of $\hat{\Omega}_\infty$ [39, 30, 29].

3.4. Calibration with Minimal Information

Step 1	Determine the catadioptric line images $\hat{\Omega}_1$ and $\hat{\Omega}_2$ using conic fitting techniques
Step 2	Compute the points $\hat{F}_{12}, \hat{B}_{12}$ where conics $\hat{\Omega}_1, \hat{\Omega}_2$ intersect and determine line $\hat{\mu}_{12} = \hat{F}_{12} \wedge \hat{B}_{12}$
Step 3	Obtain the line $\hat{\eta}_{12}$ going through the poles of $\hat{\mu}_{12}$ with respect to conics $\hat{\Omega}_1$ and $\hat{\Omega}_2$
Step 4	Compute the intersection points $\hat{M}_{12} = \hat{\mu}_{12} \wedge \hat{\pi}_\infty$ and $\hat{N}_{12} = \hat{\mu}_{12} \wedge \hat{\eta}_{12}$
Step 5	Obtain \hat{C}_{12} and \hat{N}_{12}^* such that both pairs of points $\{\hat{C}_{12}, \hat{M}_{12}\}$ and $\{\hat{N}_{12}^*, \hat{N}_{12}\}$ are harmonic conjugate with respect to $\{\hat{F}_{12}, \hat{B}_{12}\}$ ($\hat{C}_{12} = \hat{\mu}_{12} \wedge (\hat{\Omega}_i \hat{M}_{12}) = \hat{\mu}_{12} \wedge (\hat{\Omega}_j \hat{M}_{12})$ and $\hat{N}_{12}^* = \hat{\Omega}_i^* \cdot \hat{\eta}_{12} = \hat{\Omega}_j^* \cdot \hat{\eta}_{12}$).
Step 6	Given points $\hat{N}_{12}^*, \hat{N}_{12}, \hat{M}_{12}$, and \hat{C}_{12} , compute the cross-ratio $\{\hat{C}_{12}, \hat{N}_{12}; \hat{M}_{12}, \hat{O}_{12}\}$ using the relation provided by equation 3.5.
Step 7	Determine the image center \hat{O} from points $\hat{C}_{12}, \hat{N}_{12}, \hat{M}_{12}$ and the value of $\{\hat{C}_{12}, \hat{N}_{12}; \hat{M}_{12}, \hat{O}_{12}\}$
Step 8	Obtain the polar lines $\hat{\pi}_1, \hat{\pi}_2$ of the image center \hat{O} with respect to conics $\hat{\Omega}_1, \hat{\Omega}_2$ ($\hat{\pi}_1 = \hat{\Omega}_1 \hat{O}$ and $\hat{\pi}_2 = \hat{\Omega}_2 \hat{O}$).
Step 9	Determine the intersection points \hat{I}_1, \hat{J}_1 of $\hat{\pi}_1$ with $\hat{\Omega}_1$, and \hat{I}_2, \hat{J}_2 of $\hat{\pi}_2$ with $\hat{\Omega}_2$
Step 10	Estimate conic $\hat{\Omega}_\infty$ knowing that points $\hat{I}_1, \hat{J}_1, \hat{I}_2$ and \hat{J}_2 lie in the conic and that \hat{O}_1 and $\hat{\pi}_\infty$ are pole/polar with respect to $\hat{\Omega}_\infty$

Table 3.3: Calibrating an hyperbolic/elliptical system from the image of 2 lines

3. Central Catadioptric Line Projection

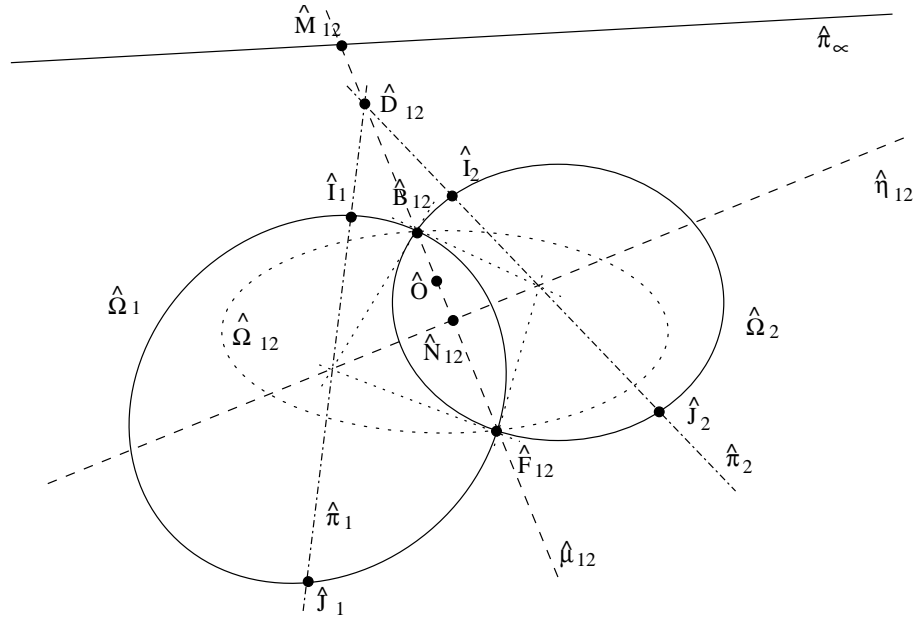


Figure 3.5: Calibrating an hyperbolic/elliptical system from the image of 2 lines

The Algorithm Steps

The algorithm steps can be followed in the scheme of Fig. 3.5. Start by obtaining the catadioptric line images $\hat{\Omega}_1$ and $\hat{\Omega}_2$. They are estimated by fitting a conic curve to points selected in the image. For conic fitting techniques see [32, 33, 76]. The next step is to determine line $\hat{\mu}_{12}$ going through the intersection points of the two conics. The line of the normals $\hat{\eta}_{12}$ is computed using the result of proposition 3.9. Lines $\hat{\mu}_{12}$, $\hat{\eta}_{12}$ have already been determined, and $\hat{\pi}_\infty$ is assumed to be known. The computation of intersection points \hat{N}_{12} , \hat{M}_{12} and the respective conjugates \hat{N}_{12}^* , \hat{C}_{12} lying on line $\hat{\mu}_{12}$ is trivial.

We have already seen that conics $\hat{\Omega}_1$ and $\hat{\Omega}_2$ define a family of catadioptric line images $\hat{\Omega}$. Conic $\hat{\Omega}$ goes through the intersection points \hat{F}_{12} , \hat{B}_{12} and the point \hat{N} associated with it, must lie on line $\hat{\eta}_{12}$. The catadioptric line image $\hat{\Omega}_{12}$, depicted in Fig. 3.5, is a particular realization of $\hat{\Omega}$. The original imaged line is on plane Π_{12} . The corresponding normal direction \bar{N}_{12} is mapped on point \hat{N}_{12} , where lines $\hat{\mu}_{12}$ and $\hat{\eta}_{12}$ intersect. Since both \hat{O} and \hat{N}_{12} lie on $\hat{\mu}_{12}$ (proposition 3.8), then this line is the locus where the major axis of $\bar{\Omega}_{12}$ is mapped (proposition 3.3). Propositions 3.4, 3.5 and equation 3.5 hold, and the image center can be

3.4. Calibration with Minimal Information

computed as described on steps 6 and 7.

The polar lines of the image center \hat{O} are $\hat{\pi}_1$ and $\hat{\pi}_2$. The computation of the intersection points $\hat{I}_1, \hat{J}_1, \hat{I}_2$ and \hat{J}_2 is trivial. From proposition 3.2 comes that these four points must lie on conic $\hat{\Omega}_\infty$. Moreover \hat{O} and $\hat{\pi}_\infty$ are pole/polar with respect to $\hat{\Omega}_\infty$ (corollary 3.2). We have the necessary five constraints to determine the absolute conic locus.

Additional Remarks

The proposed algorithm is not valid for central parabolic images. Accordingly to proposition 3.7, if the system is parabolic ($\xi=1$), then points $\hat{N}_{12}, \hat{M}_{12}$ are coincident with points $\hat{C}_{12}, \hat{N}_{12}^*$. The cross-ratios $\{\hat{N}_{12}^*, \hat{N}_{12}; \hat{M}_{12}, \hat{C}_{12}\}$ and $\{\hat{C}_{12}, \hat{N}_{12}; \hat{M}_{12}, \hat{O}_{12}\}$ become respectively 0 and 1, despite the imaged lines. Under these circumstances the image center \hat{O} can not be recovered using the result of equation 3.5.

The hyperbolic/elliptical system can be calibrated from the image of two lines. However it seems that more information is needed to obtain the calibration for the parabolic situation. This supports the conjecture advanced in [34]. According to the conjecture, the minimum number of lines to calibrate an hyperbolic/elliptical system is two, while to calibrate a parabolic system is three. The explanation has to do with the number of unknowns and constraints. The goal is to estimate the image of the absolute conic $\hat{\Omega}_\infty$, which has five unknown parameters. Since each line introduces two additional unknowns, then using two line images we have a total of nine parameters to determine. The lines are imaged in conic curves by an hyperbolic/elliptical system. Each conic has five degrees of freedom, thus two line images provide ten constraints. Since the constraints are more than the unknowns the system can be calibrated. Consider the statement made on proposition 3.6 for parabolic systems. Any paracatadioptric line image $\hat{\Omega}$ must go through the locus of the circular points and verify the relation $\hat{G}^t \hat{\Omega} \hat{H} = 0$ where \hat{G} and \hat{H} are two fixed points. Thus, the first line image provides five constraints, while any additional line only gives two constrains more. The calibration from two line images is an underdetermined problem, since there are only seven constraints for nine unknowns. We will return to this discussion later on.

3. Central Catadioptric Line Projection

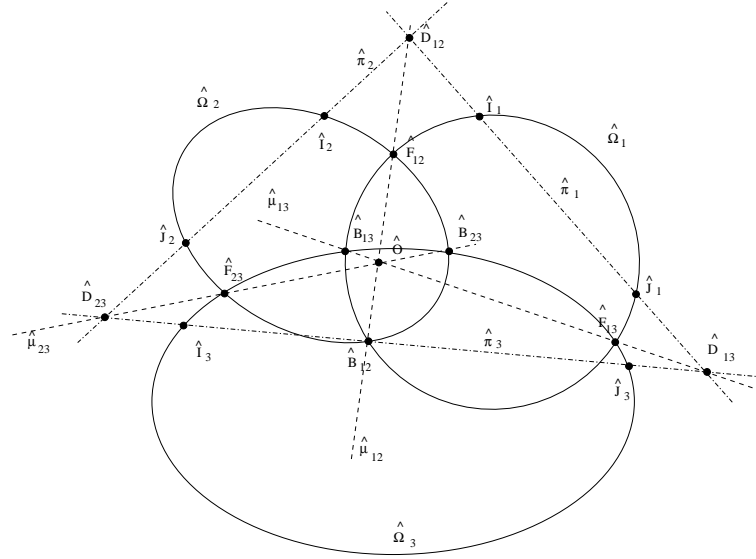


Figure 3.6: Calibrating a central catadioptric system from the image of 3 lines (Part I)

3.4.2 Full Calibration of a Central Catadioptric System from Three Lines

It is possible to calibrate any central catadioptric system from the image of three lines (the conventional perspective camera is not included). The lines must be projected into three non-degenerate conics that do not intersect in the same pair of points. The imaged lines can not be coplanar with the mirror symmetry axis, nor lie in planes that intersect in a same line going through the effective viewpoint \hat{O} (Fig. 3.4). The proposed algorithm is summarized on Tab. 3.4. It is shown that three lines are enough to obtain the locus of the absolute conic $\hat{\Omega}_\infty$, the position of line $\hat{\pi}_\infty$ and the parameter ξ without further information.

The Algorithm Steps

Consider the scheme of Fig. 3.6. The three lines are projected on conics $\hat{\Omega}_1, \hat{\Omega}_2$ and $\hat{\Omega}_3$, estimated using a conic fitting technique [32, 33, 76]. Each pair of conics $\hat{\Omega}_i, \hat{\Omega}_j$ intersect on two real points which define a line $\hat{\mu}_{ij}$ ($ij = 12, 13, 23$). The three lines $\hat{\mu}_{12}, \hat{\mu}_{13}$ and $\hat{\mu}_{12}$ must intersect in the image center \hat{O} (proposition

3.4. Calibration with Minimal Information

Step 1	Determine the catadioptric line images $\hat{\Omega}_1$, $\hat{\Omega}_2$ and $\hat{\Omega}_3$ using conic fitting techniques
Step 2	For each pair of conics $\hat{\Omega}_i$, $\hat{\Omega}_j$, compute the intersection points \hat{F}_{ij} , \hat{B}_{ij} and determine line $\hat{\mu}_{ij} = \hat{F}_{ij} \wedge \hat{B}_{ij}$ ($ij = 12, 13, 23$)
Step 3	Obtain the image center \hat{O} as the intersection point of lines $\hat{\mu}_{12}$, $\hat{\mu}_{13}$ and $\hat{\mu}_{23}$.
Step 4	For each conic $\hat{\Omega}_i$ compute the polar line $\hat{\pi}_i$ of the image center \hat{O} ($i = 1, 2, 3$).
Step 5	For each conic curve obtain the points \hat{I}_i and \hat{J}_i where line $\hat{\pi}_i$ intersects $\hat{\Omega}_i$ ($i = 1, 2, 3$)
Step 6	Estimate conic $\hat{\Omega}_\infty$ going through points $\hat{I}_1, \hat{J}_1, \hat{I}_2, \hat{J}_2, \hat{I}_3$ and \hat{J}_3
Step 7	For each pair of conics $\hat{\Omega}_i$, $\hat{\Omega}_j$ determine line $\hat{\eta}_{ij}$ going through the poles of $\hat{\mu}_{ij}$ with respect to $\hat{\Omega}_i, \hat{\Omega}_j$ ($ij = 12, 13, 23$)
Step 8	Obtain the intersection points $\hat{N}_1 = \hat{\eta}_{12} \wedge \hat{\eta}_{13}$, $\hat{N}_2 = \hat{\eta}_{12} \wedge \hat{\eta}_{23}$ and $\hat{N}_3 = \hat{\eta}_{13} \wedge \hat{\eta}_{23}$.
Step 9	Determine the line $\hat{\mu}_i = \hat{N}_i \wedge \hat{O}$ associated with the catadioptric line image $\hat{\Omega}_i$ ($i = 1, 2, 3$)
Step 10	Estimate the locus of the line at infinity $\hat{\pi}_\infty$ knowing that it goes through the poles of $\hat{\mu}_i$ with respect to conic $\hat{\Omega}_i$ ($i = 1, 2, 3$)
Step 11	For each conic $\hat{\Omega}_i$ determine the intersection point $\hat{D}_i = \hat{\mu}_i \wedge \hat{\pi}_i$ and the pole $\hat{C}_i = \hat{\Omega}_i^* \cdot \hat{\pi}_\infty$ which must lie on line $\hat{\mu}_i$ ($i = 1, 2, 3$)
Step 12	The mirror parameter ξ is provided by $\sqrt{\{\hat{O}, \hat{D}_i; \hat{N}_i, \hat{C}_i\}}$ ($i = 1, 2, 3$)

Table 3.4: Calibrating a central catadioptric system using 3 lines

3. Central Catadioptric Line Projection

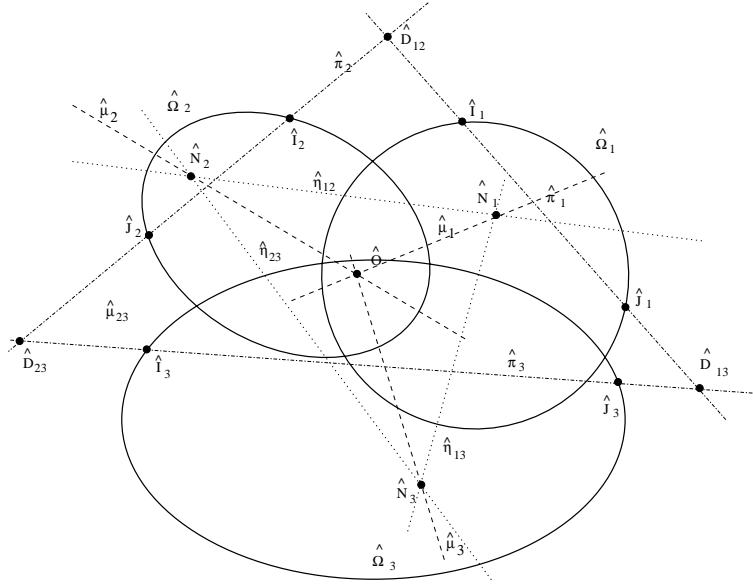


Figure 3.7: Calibrating a central catadioptric system from the image of 3 lines (Part II)

3.8). Given the image center compute the polar line with respect to each conic image. The polar line $\hat{\pi}_i$ intersects the corresponding conic $\hat{\Omega}_i$ on two points \hat{I}_i, \hat{J}_i ($i = 1, 2, 3$). Accordingly to proposition 3.2 these points lie on the conic $\hat{\Omega}_\infty$, which is the locus where the absolute conic is mapped by collineation H_c . Conic $\hat{\Omega}_\infty$ is estimated using the six points $\hat{I}_1, \hat{J}_1, \hat{I}_2, \hat{J}_2, \hat{I}_3, \hat{J}_3$.

Each pair of imaged lines has a line of normals associated with it, which is mapped on $\hat{\eta}_{ij}$ in the catadioptric image plane ($ij = 12, 13, 23$). The line goes through the poles of $\hat{\mu}_{ij}$ with respect to conics $\hat{\Omega}_i$ and $\hat{\Omega}_j$ (proposition 3.9). Remember that both points \hat{N}_i and \hat{N}_j lie on $\hat{\eta}_{ij}$. Thus the pairs of lines $\hat{\eta}_{12}, \hat{\eta}_{13}; \hat{\eta}_{12}, \hat{\eta}_{23}$ and $\hat{\eta}_{13}, \hat{\eta}_{23}$ intersect on points \hat{N}_1, \hat{N}_2 and \hat{N}_3 (Fig. 3.7). The determination of line $\hat{\mu}_i$ associated to the line image $\hat{\Omega}_i$ knowing both \hat{O} and \hat{N}_i is trivial (proposition 3.3). The locus of the line at infinity can be easily determined taking into account that the poles of $\hat{\mu}_1, \hat{\mu}_2$ and $\hat{\mu}_3$ must lie in $\hat{\pi}_\infty$ (corollary 3.1). The ξ parameter of the system mirror is computed using the result of proposition 3.5.

3.5 Closure

The model established for central catadioptric sensors and the geometric properties derived can be used in many different ways. A possible application is 3D reconstruction from panoramic images as described in [66]. It is possible to directly recover the orientation of any world plane Π from the catadioptric image of two sets of parallel lines lying on it. Consider that the directions of the two sets of lines are respectively $\bar{\mathbf{D}}$ and $\bar{\mathbf{D}}'$. The direction points are mapped in the catadioptric image plane at points $\hat{\mathbf{D}}$ and $\hat{\mathbf{D}}'$. These points can be easily determined from the line images using the result of corollary 3.5. Since both $\bar{\mathbf{D}}$ and $\bar{\mathbf{D}}'$ must lie on the horizon line of plane Π then π is mapped on $\hat{\pi} = \hat{\mathbf{D}} \wedge \hat{\mathbf{D}}'$. If \mathbf{H}_c is known then the horizon line of Π is $\pi = \mathbf{H}_c^t \hat{\pi}$ and the orientation of the plane is recovered.

The derived theory can also be used for calibration purposes as discussed on section 3.4.1 and 3.4.2. The algorithms summarized on Tab. 3.3 and 3.4 show that two line images are enough to calibrate an hyperbolic/elliptical sensor and any central panoramic system can be fully calibrated from the image of three lines. If the catadioptric lines images are correctly determined then the calibration is straightforward. However the estimation of these conics using image points is hard to accomplish. There are several algorithms to fit a conic curve to data points. In [32, 76] some of these algorithms are reviewed and their performance is evaluated. A robust conic fitting algorithm has to cope with noisy data points, biasing due to curvature and partial occlusion. The occlusion problem is of particular importance for our purposes. By occlusion we mean that the available data points lie on a small arc of the curve. In these cases, it is very hard to estimate the correct conic curve, even for small amounts of noise.

Fig. 3.8 is the image of a white board acquired with an hyperbolic sensor. An edge detector is then used to obtain the sides of the board. Since the board is a square, the corresponding sides are catadioptric line images. The data points, picked on each side of the board, are fitted by a conic curve using the approximate mean square algorithm [32]. As it can be observed the estimated curves go through the sides of the square. According to proposition 3.8, each pair of catadioptric line images must intersect on two real points. These points define a line going through the image center. Since there is a single image center, all the lines going through the intersection points must intersect in a single point. The

3. Central Catadioptric Line Projection



Figure 3.8: Estimating catadioptric line images using standard conic fitting techniques

estimated conic curves do not satisfy this property. The catadioptric line images are not correctly determined due to occlusion. The next chapter discusses the application of the proposed theory in solving the problem of correctly estimate line images in the catadioptric image plane.

Chapter 4

Using Line Images for Calibration Purposes

The previous chapter proves that an hyperbolic/elliptical system can be calibrated from the image of two lines and that any central catadioptric sensor can be fully calibrated using a minimum number of three line images.

In general lines in the scene are mapped into conic loci in the catadioptric image plane. The calibration procedure requires the computation of the intersection points between lines and conic and between pairs of conics. Section 4.1 proposes algorithms to accomplish these tasks. It is shown that the problem of intersecting a line and a conic has a closed form solution. A method to compute the intersection between two conics is also proposed. Section 4.2 focuses on the estimation of conic curves. Five conic fitting methods are compared and discussed. It is shown that in general none of these conventional techniques copes well with occlusion. Section 4.3 uses the geometric properties of catadioptric line projection to increase the robustness of the conic fitting. The estimated set of conic curves is used to calibrate the system following the steps summarized in Tab. 3.3 and 3.4. The calibration results are evaluated using both real images and simulated data.

4.1 Working with Conic Curves

The present section presents a closed form solution to compute the intersection of a line with a conic curve, and a numerically stable method to determine the

4. Using Line Images for Calibration Purposes

intersection of two conics. We start by reviewing the concepts of conic locus, conic envelope and skew symmetric matrix. For a detailed study on these subjects see [59, 38].

A conic locus is an algebraic locus of second order. It can be parameterized by a 3×3 symmetric matrix Ω (equation 4.1). Consider a point \mathbf{x} in the projective plane \mathbf{P}^2 . The point lies in the conic curve Ω if, and only if, $\mathbf{x}^t \Omega \mathbf{x} = 0$.

$$\Omega = \begin{bmatrix} a & b & d \\ b & c & e \\ d & e & f \end{bmatrix} \quad (4.1)$$

The conic locus is said to be proper or degenerate according as it is or is not irreducible. For the former the corresponding matrix Ω is full rank, for the latter the matrix is rank deficient. There are two distinct kinds of degenerate conic locus: a pair of distinct lines (rank 2 matrix), and a repeated line (rank 1 matrix).

The dual figure of the conic locus Ω is the conic envelope Ω^* provided in equation 4.2. Notice that if Ω is a full rank matrix, then $\Omega^* = \det(\Omega) \Omega^{-1}$.

$$\Omega^* = \begin{bmatrix} fc - e^2 & ed - bf & be - cd \\ ed - bf & af - d^2 & bd - ae \\ be - cd & bd - ae & ac - b^2 \end{bmatrix} \quad (4.2)$$

Consider the 3×1 vectors $\mathbf{v} = (v_x, v_y, v_z)^t$ and \mathbf{w} . We can always associate a 3×3 skew symmetric matrix $\tilde{\mathbf{v}}$ with vector \mathbf{v} (equation 4.3). Since matrix $\tilde{\mathbf{v}}$ is skew symmetric then it verifies $\tilde{\mathbf{v}}^t = -\tilde{\mathbf{v}}$. The vector product of \mathbf{v} and \mathbf{w} is $\mathbf{v} \wedge \mathbf{w} = \tilde{\mathbf{v}} \mathbf{w}$.

$$\tilde{\mathbf{v}} = \begin{bmatrix} 0 & -v_z & v_y \\ v_z & 0 & -v_x \\ -v_y & v_x & 0 \end{bmatrix} \quad (4.3)$$

4.1.1 Intersection of a Line with a Conic

Fig. 4.1 depicts a line $\mathbf{r} = (r_x, r_y, r_z)^t$ and a conic curve Ω . A line intersects a conic locus in two points. We aim to determine points \mathbf{P}^+ and \mathbf{P}^- where \mathbf{r} and Ω intersect. Consider point $\mathbf{P}_1 = (-r_z, -r_z, r_x + r_y)^t$ lying on line \mathbf{r} ($\mathbf{P}_1^t \mathbf{r} = 0$). Notice that $\mathbf{P}_1 = \mathbf{I}_s \mathbf{r}$ with \mathbf{I}_s the 3×3 matrix provided by equation 4.4.

4.1. Working with Conic Curves

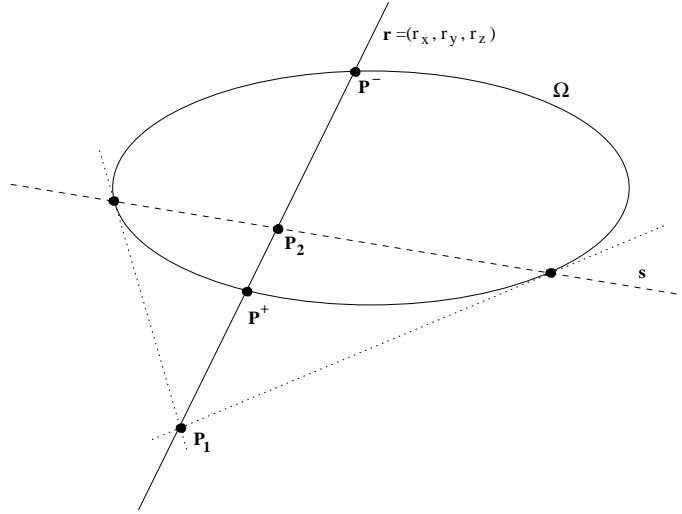


Figure 4.1: Intersecting a line with a conic

$$\mathbf{I}_s = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & -1 \\ 1 & 1 & 0 \end{bmatrix} \quad (4.4)$$

The polar line of \mathbf{P}_1 with respect to conic Ω is $s = \Omega\mathbf{P}_1$. Lines r and s intersect on point $\mathbf{P}_2 = r \wedge s = \tilde{r}\Omega\mathbf{P}_1$. If \mathbf{P}_1 and \mathbf{P}_2 are distinct then any point \mathbf{P} , lying on line r , can be written as a linear combination of the two conjugate points (equation 4.5).

$$\mathbf{P}(\lambda) = \mathbf{P}_2 + \lambda\mathbf{P}_1 \quad (4.5)$$

Point $\mathbf{P}(\lambda)$ lies in the line r for any value of λ . We aim to find the points $\mathbf{P}(\lambda)$ belonging to the conic curve Ω . Thus,

$$\mathbf{P}(\lambda)^t \cdot \Omega \cdot \mathbf{P}(\lambda) = 0$$

Replacing $\mathbf{P}(\lambda)$ by the result of equation 4.5 and taking into account that \mathbf{P}_1 and \mathbf{P}_2 are conjugate with respect to the conic curve Ω yields

$$\mathbf{P}_1^t \Omega \mathbf{P}_1 \lambda^2 + \mathbf{P}_2^t \Omega \mathbf{P}_2 = 0$$

4. Using Line Images for Calibration Purposes

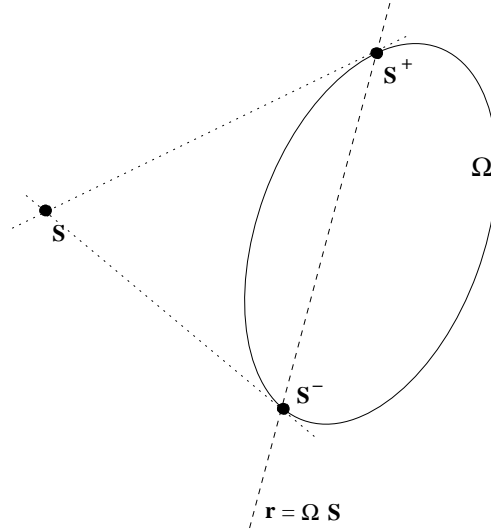


Figure 4.2: The tangents to the conic going through a point

Solving the above equation in order to λ arises

$$\begin{aligned} \lambda &= \pm \sqrt{-\frac{\mathbf{P}_2^t \Omega \mathbf{P}_2}{\mathbf{P}_1^t \Omega \mathbf{P}_1}} \\ &= \pm \sqrt{-\mathbf{r}^t \Omega^* \mathbf{r}} \end{aligned}$$

The intersection points \mathbf{P}^+ and \mathbf{P}^- are determined by replacing λ in equation 4.5. Making $\mathbf{P}_1 = \mathbf{I}_s \mathbf{r}$ and $\mathbf{P}_2 = \tilde{\mathbf{r}} \Omega \mathbf{I}_s \mathbf{r}$ yields

$$\mathbf{P}^\pm = (\pm \sqrt{-\mathbf{r}^t \Omega^* \mathbf{r}} \mathbf{I} + \tilde{\mathbf{r}} \Omega) \mathbf{I}_s \mathbf{r} \quad (4.8)$$

The formula provided in 4.8 computes the points \mathbf{P}^+ , \mathbf{P}^- where line \mathbf{r} intersects the conic curve Ω . The formula holds even for degenerate conics. If matrix Ω has rank 2 then the conic locus is a pair of distinct lines which intersect \mathbf{r} in two distinct points. If Ω has rank 1 then the conic is a repeated line \mathbf{m} . Since $\Omega = \mathbf{m} \mathbf{m}^t$ then $\Omega^* = 0$ (equation 4.2) and $\mathbf{P}^+ = \mathbf{P}^- = \tilde{\mathbf{r}} \mathbf{m}$ (equation 4.8). The formula is still valid for a line \mathbf{r} tangent to the conic locus Ω . Since the line \mathbf{r} lies in the conic envelope Ω^* then $\mathbf{r}^t \Omega^* \mathbf{r} = 0$. Points \mathbf{P}^+ and \mathbf{P}^- are coincident and correspond to the tangency point.

4.1. Working with Conic Curves

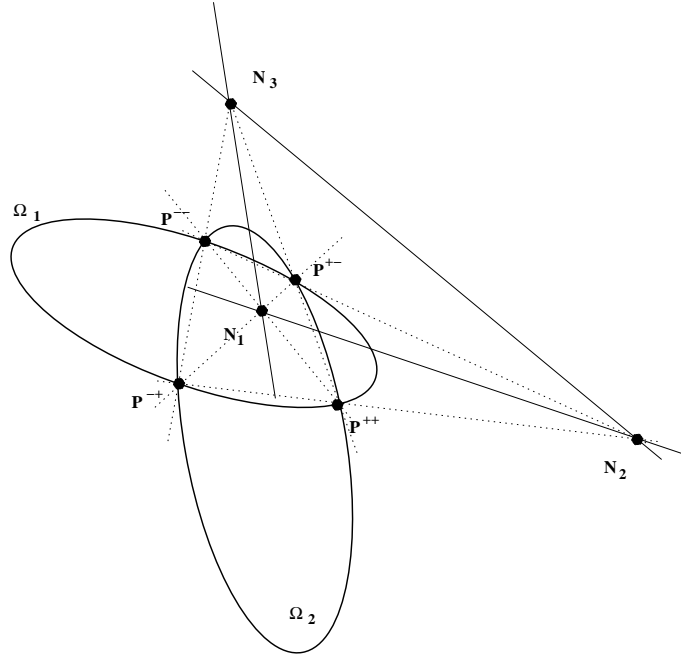


Figure 4.3: Intersection of two conics

Consider a generic point S in the projective plane. The result of equation 4.8 can be used to compute the pair of conjugate points S^+, S^- lying in the conic section (Fig. 4.2). The lines tangent to the conic Ω on points S^+ and S^- must intersect in S . Since points S^+, S^- are conjugate to S , then they must lie in the corresponding polar line $r = \Omega S$. Points S^+, S^- are the loci where line r and conic Ω intersect. Replacing r by ΩS in equation 4.8 yields

$$S^\pm = (\pm \sqrt{-S^t \Omega (\Omega^*) \Omega S} I + (\tilde{\Omega} S) \Omega) I_s \Omega S$$

Assuming that the conic is proper and matrix Ω is full rank comes that $(\tilde{\Omega} S) = \det(\Omega) \Omega^{-t} \tilde{S} \Omega^{-1}$. Moreover Ω is a symmetric matrix and $\det(\Omega) \Omega^{-1} = \Omega^*$. After some algebraic manipulation results

$$S^\pm = (\pm \sqrt{-\det(\Omega) S^t \Omega S} I + \Omega^* \tilde{S}) I_s \Omega S \quad (4.10)$$

4.1.2 Intersection of Two Conic Curves

Consider the conic curves Ω_1 and Ω_2 depicted in Fig. 4.3. Two conic curves intersect in four points \mathbf{P}^{++} , \mathbf{P}^{+-} , \mathbf{P}^{-+} and \mathbf{P}^{--} . This section derives an algorithm to compute the intersection points of a pair of conics. The proposed algorithm is numerically stable and computationally efficient.

The Self-Polar Triangle Common to a Pair of Conic Curves

Equation 4.11 defines the pencil of conic curves $\Omega(\lambda)$. The pencil of conics is the infinite set of conic curves $\Omega(\lambda)$ which are a linear combination of the base conics Ω_1 and Ω_2 . Notice that if \mathbf{P} is a point common to both Ω_1 and Ω_2 ($\mathbf{P}^t \Omega_1 \mathbf{P} = 0$ and $\mathbf{P}^t \Omega_2 \mathbf{P} = 0$), then \mathbf{P} lies in $\Omega(\lambda)$ for any value of λ ($\mathbf{P}^t \Omega(\lambda) \mathbf{P} = 0$). A pair of conic curves always intersect in four points which can be real or complex, distinct or coincident. The pencil of conics determined by Ω_1 and Ω_2 is simply the system of all conics through the four common points of Ω_1 and Ω_2

$$\Omega(\lambda) = \Omega_1 + \lambda \Omega_2 \quad (4.11)$$

We shall assume from now on, except when there is an explicit statement to the contrary, that the four common points of the pencil $\Omega(\lambda)$ are distinct. Points \mathbf{P}^{++} , \mathbf{P}^{+-} , \mathbf{P}^{-+} , \mathbf{P}^{--} define a quadrangle. The diagonal triangle of the quadrangle is depicted in Fig. 4.3. The vertices of the triangle are \mathbf{N}_1 , \mathbf{N}_2 and \mathbf{N}_3 . It can be shown that triangle $\mathbf{N}_1 \mathbf{N}_2 \mathbf{N}_3$ is the single common self-polar triangle for all the conics in the pencil [59]. This means that the pair of vertices $\mathbf{N}_1, \mathbf{N}_2$; $\mathbf{N}_1, \mathbf{N}_3$ and $\mathbf{N}_2, \mathbf{N}_3$ are conjugate with respect to all conics $\Omega(\lambda)$. Putting it in another way comes that $\mathbf{N}_1^t \cdot \Omega(\lambda) \cdot \mathbf{N}_2 = 0$, $\mathbf{N}_1^t \cdot \Omega(\lambda) \cdot \mathbf{N}_3 = 0$ and $\mathbf{N}_2^t \cdot \Omega(\lambda) \cdot \mathbf{N}_3 = 0$ for any value of λ .

In general four points define three distinct pairs of lines. Since the common self-polar triangle is diagonal to the quadrangle defined by the points where the base conics intersect, arises that the three vertices \mathbf{N}_1 , \mathbf{N}_2 and \mathbf{N}_3 are the intersection points of the three pairs of lines defined by points \mathbf{P}^{++} , \mathbf{P}^{+-} , \mathbf{P}^{-+} and \mathbf{P}^{--} (Fig. 4.3). Moreover each one of these pairs of lines is a degenerate conic of the pencil defined in equation 4.11. Consider the following third order equation in λ

4.1. Working with Conic Curves

$$\det(\Omega_1 + \lambda\Omega_2) = 0$$

It can be shown that if conics Ω_1 and Ω_2 intersect in four distinct points then there are three distinct λ solutions. Assume that the roots are λ_a , λ_b and λ_c . Replacing in equation 4.11 we obtain matrices Ω_a , Ω_b and Ω_c . Each matrix has rank 2 and corresponds to a degenerate conic (a pair of lines). Since the degenerate conics belong to the pencil $\Omega(\lambda)$ then they are the three line pairs going through the intersections of Ω_1 and the Ω_2 . Points N_1 , N_2 and N_3 can be determined by computing the null spaces of matrices Ω_a , Ω_b and Ω_c .

Computing the Intersection Points Using the Common Self-Polar Triangle

Points N_1 , N_2 , N_3 are the vertices of the self-polar triangle associated with the pencil of conics $\Omega(\lambda)$ (equation 4.11). We proved that, knowing a pair of conics of the pencil, it is possible to determine these vertices by solving a third order equation and performing some additional algebraic manipulation. If the elements of pencil $\Omega(\lambda)$ intersect into four distinct points, then the vertices N_1 , N_2 , N_3 are non collinear. Any point P in the projective plane can be written as a linear combination of three non collinear points. Thus, fixating the scale factor and using $[N_1N_2N_3]$ as a basis arises

$$P(\phi, \theta) = N_1 + \phi N_2 + \theta N_3 \quad (4.13)$$

We aim to determine the parameters ϕ and θ such that the corresponding point $P(\phi, \theta)$ provided by equation 4.13 lies in both base conics Ω_1 and Ω_2 . The equation to be solved is

$$\begin{cases} P(\phi, \theta)^t \cdot \Omega_1 \cdot P(\phi, \theta) = 0 \\ P(\phi, \theta)^t \cdot \Omega_2 \cdot P(\phi, \theta) = 0 \end{cases}$$

Replacing $P(\phi, \theta)$ by the result of equation 4.13 and taking into account that N_1 , N_2 and N_3 are conjugate points with respect to both conics yields

$$\begin{cases} N_1^t \Omega_1 N_1 + \phi^2 N_2^t \Omega_1 N_2 + \theta^2 N_3^t \Omega_1 N_3 = 0 \\ N_1^t \Omega_2 N_1 + \phi^2 N_2^t \Omega_2 N_2 + \theta^2 N_3^t \Omega_2 N_3 = 0 \end{cases}$$

4. Using Line Images for Calibration Purposes

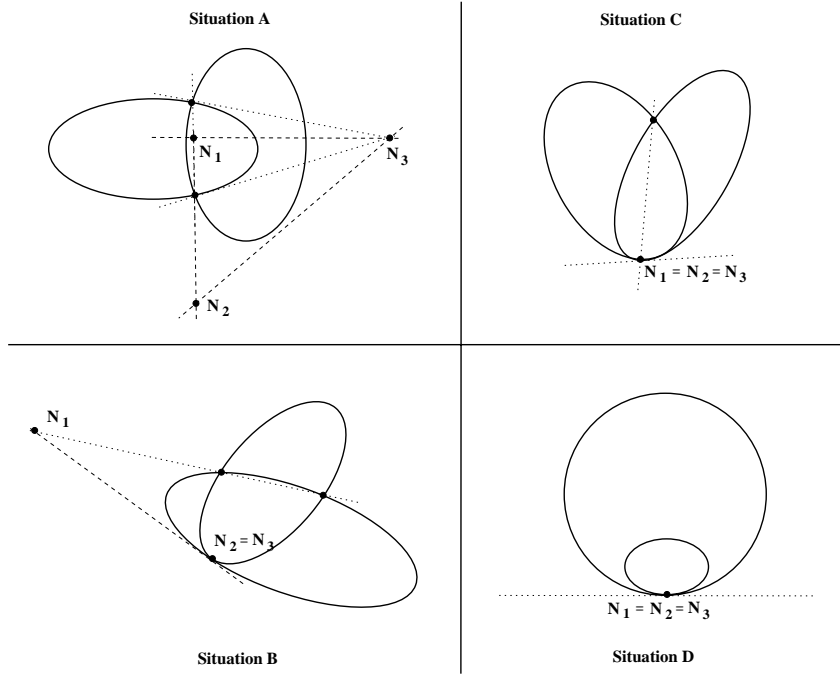


Figure 4.4: Degenerate Configurations

Solving the system of equations in order to ϕ and θ results

$$\begin{cases} \phi^\pm = \pm \sqrt{\frac{N_1^t \Omega_1 N_1 N_3^t \Omega_2 N_3 - N_1^t \Omega_2 N_1 N_3^t \Omega_1 N_3}{N_3^t \Omega_1 N_3 N_2^t \Omega_2 N_2 - N_3^t \Omega_2 N_3 N_2^t \Omega_1 N_2}} \\ \theta^\pm = \pm \sqrt{\frac{N_2^t \Omega_1 N_2 N_1^t \Omega_2 N_1 - N_2^t \Omega_2 N_2 N_1^t \Omega_1 N_1}{N_3^t \Omega_1 N_3 N_2^t \Omega_2 N_2 - N_3^t \Omega_2 N_3 N_2^t \Omega_1 N_2}} \end{cases} \quad (4.16)$$

The points where conics Ω_1 and Ω_2 intersect can be determined by replacing ϕ and θ in equation 4.13. The intersection points are $\mathbf{P}^{++} = \mathbf{P}(\phi^+, \theta^+)$, $\mathbf{P}^{+-} = \mathbf{P}(\phi^+, \theta^-)$, $\mathbf{P}^{-+} = \mathbf{P}(\phi^-, \theta^+)$ and $\mathbf{P}^{--} = \mathbf{P}(\phi^-, \theta^-)$ with ϕ^+ , ϕ^- , θ^+ and θ^- provided in equation 4.16.

Degenerate Configurations

We have just presented an algorithm to determine the points where a pair of conics curves intersect. The idea is to use the vertices of the self polar triangle common to both conics to reference the points in the projective plane. A generic point \mathbf{P} is written as a linear combination of points \mathbf{N}_1 , \mathbf{N}_2 and \mathbf{N}_3 (equation 4.13). Since the base points are conjugate with respect to both conics, the system of equations

4.2. Fitting a Conic Curve to Image Points

can be simplified and an explicit solution can be easily found (equation 4.16).

The approach is general and can be applied whenever the conics have a common self polar triangle. For this situation the vertices are non collinear and $\mathbf{N}_1\mathbf{N}_2\mathbf{N}_3$ can be used as a basis to reference the other points in the projective plane (equation 4.13). However there are a couple of configurations for which the self polar triangle degenerates and the vertices become collinear. In this case it is not possible to establish the desired basis and the algorithm can not be used.

Assume that conics Ω_1 and Ω_2 intersect in four distinct points (Fig. 4.3). It can be shown that no three of these points can be collinear. The corresponding quadrangle is never degenerate, nor the diagonal triangle associated with it. Since the vertices are not collinear then there is a valid basis $\mathbf{N}_1\mathbf{N}_2\mathbf{N}_3$ and the algorithm can be applied to determine the four intersection points.

Fig. 4.4 depicts configurations for which the four intersection points are not distinct. Consider situation A where the pair of conics intersect in two pairs of coincident points. Despite of the fact that the quadrangle formed by the conic intersections degenerates into a line segment, the common self polar triangle does not degenerate. Points \mathbf{N}_1 , \mathbf{N}_2 and \mathbf{N}_3 are not collinear and the algorithm can be used. In situation B only two of the intersection points are coincident. The associated quadrangle degenerates in a triangle and the self-polar triangle degenerates in a line segment. Since points \mathbf{N}_2 , \mathbf{N}_3 are coincident we can not establish a reference basis as done in equation 4.13. The proposed approach can not be applied to this particular configuration. The same happens for situations C and D. In the former the conics intersect in three coincident points and in the latter the four intersection points are the same. In both case the self-polar triangle degenerates into a point. Points \mathbf{N}_1 , \mathbf{N}_2 and \mathbf{N}_3 are coincident and the desired basis can not be established.

4.2 Fitting a Conic Curve to Image Points

In last section the conic loci were parameterized by a 3×3 symmetric matrix Ω (equation 4.17). Since the conic curve in the projective plane \mathbf{P}^2 has 5 independent degrees of freedom (DOF), it is natural to represent such locus by a point in \mathbf{P}^5 [59]. An alternative parameterization for the conic Ω is provided in 4.17. The representation of conic loci using points of \mathbf{P}^5 can be extremely useful and con-

4. Using Line Images for Calibration Purposes

venient. For now on we will assume both parameterizations without distinction.

$$\omega = (a, b, c, d, e, f)^t \quad (4.17)$$

A conic fitting algorithm determines the conic curve that best fits the data points according to a certain distance metric. There are several approaches to estimate a conic curve using points [76, 32]. They differ between each other by the criteria, or metric, that is minimized in the fitting process. This section reviews some well known conic fitting methods and compares their performance.

4.2.1 Least Square Fitting Based on Algebraic Distances

Consider the set of distinct points $\mathbf{x}_i = (x_i, y_i)^t$ with $i = 1, 2 \dots M$ lying on a plane ($M \geq 5$). The algebraic distance α_i between a point \mathbf{x}_i and a certain conic locus ω is given by equation 4.18. If the point lies in the curve then the corresponding algebraic distance α_i must be null.

$$\alpha_i = \alpha(\mathbf{x}_i) = ax_i^2 + 2bx_iy_i + cy_i^2 + 2dx_i + 2ey_i + f, i = 1 \dots M \quad (4.18)$$

The $M \times 6$ matrix \mathbf{A} provided in equation 4.19 is called the design matrix. The vector of the algebraic distances between the conic curve and the entire data set is $(\alpha_1, \alpha_2, \dots, \alpha_M)^t = \mathbf{A}\omega$. It is straightforward that if the set of points lie in the conic curve then \mathbf{A} is rank 5 and ω is in the null space of the matrix.

$$\mathbf{A} = \begin{bmatrix} x_1^2 & 2x_1y_1 & y_1^2 & 2x_1 & 2y_1 & 1 \\ x_2^2 & 2x_2y_2 & y_2^2 & 2x_2 & 2y_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_M^2 & 2x_My_M & y_M^2 & 2x_M & 2y_M & 1 \end{bmatrix} \quad (4.19)$$

Assume that the set of data points lying in the conic has more than 5 elements. In general the points are corrupted with noise and matrix \mathbf{A} is rank 6. If the design matrix is full rank then there is no null space and the unique solution of equation $\mathbf{A}\omega = 0$ is the trivial one ($\omega = 0$). A common practice is to fit the data points by the conic which minimizes the sum of the square of the algebraic distances [19]. Thus we aim to find a minima of the function ϕ provided in equation 4.20.

4.2. Fitting a Conic Curve to Image Points

$$\phi(\omega) = \sum_{i=1}^M \alpha_i^2 = \omega^t \mathbf{A}^t \mathbf{A} \omega \quad (4.20)$$

Clearly the trivial solution $\omega = 0$ is a global minima of ϕ . In order to avoid it, we must constrain ω . Several different methods are proposed in the literature [55, 69, 32]. We focus exclusively on direct methods for which the fitting problem can be solved naturally by an eigensystem.

Normal Least Squares (LMS)

The LMS method estimates the conic curve ω that minimizes the algebraic distance ϕ under the constraint $\hat{\omega}^t \hat{\omega} = 1$. The objective function is provided below where the constraint is introduced using a Lagrange multiplier λ

$$\phi_{lms}(\omega, \lambda) = \omega^t \mathbf{A}^t \mathbf{A} \omega + \lambda(\omega^t \omega - 1)$$

The conic curve ω that minimizes ϕ_{lms} is determined by solving the eigensystem $\mathbf{A}^t \mathbf{A} \hat{\omega} = \lambda \hat{\omega}$. The minimizer is the eigenvector corresponding to the smallest eigenvalue of matrix $\mathbf{A}^t \mathbf{A}$ [55].

Approximate Mean Squares (AMS)

The approximate mean square metric has been introduced by Taubin in [69]. The proposed conic fitting method minimizes the algebraic distance under the constraint $\omega^t (\mathbf{A}_x^t \mathbf{A}_x + \mathbf{A}_y^t \mathbf{A}_y) \omega = 1$ where \mathbf{A}_x and \mathbf{A}_y are the partial derivatives of \mathbf{A} .

$$\mathbf{A}_x = \begin{bmatrix} 2x_1 & 2y_1 & 0 & 2 & 0 & 0 \\ 2x_2 & 2y_2 & 0 & 2 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 2x_M & 2y_M & 0 & 2 & 0 & 0 \end{bmatrix}; \mathbf{A}_y = \begin{bmatrix} 0 & 2x_1 & 2y_1 & 0 & 2 & 0 \\ 0 & 2x_2 & 2y_2 & 0 & 2 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 2x_M & 2y_M & 0 & 2 & 0 \end{bmatrix}$$

Introducing the constraint using a Lagrange multiplier yields

$$\phi_{ams}(\omega, \lambda) = \omega^t \mathbf{A}^t \mathbf{A} \omega + \lambda(\omega^t (\mathbf{A}_x^t \mathbf{A}_x + \mathbf{A}_y^t \mathbf{A}_y) \omega - 1)$$

4. Using Line Images for Calibration Purposes

The solution ω which minimizes the final objective function ϕ_{ams} can be analytically determined by solving the generalized eigensystem $\mathbf{A}^t \mathbf{A} \hat{\omega} = \lambda(\mathbf{A}_x^t \mathbf{A}_x + \mathbf{A}_y^t \mathbf{A}_y) \hat{\omega}$ [69]. The conic curve estimation is provided by the eigenvector corresponding to the smallest eigenvalue.

Direct Least Square Fitting of Ellipses (FF)

The estimation method proposed by Fitzgibbon and Fisher in [33] is ellipse specific. Consider the following 6×6 matrix

$$\mathbf{C} = \begin{bmatrix} 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

If a conic ω verifies $\omega^t \mathbf{C} \omega = 1$ then it must be a circle/ellipse. The conic fitting method proposed in [33] estimates the curve by minimizing the algebraic distance to the data points under the constraint $\omega^t \mathbf{C} \omega = 1$. The resultant objective function is provided below where the constraint is introduced using a Lagrange multiplier

$$\phi_{ff}(\omega, \lambda) = \omega^t \mathbf{A}^t \mathbf{A} \omega + \lambda(\omega^t \mathbf{C} \omega - 1)$$

It can be proved that the ellipse/circle ω which minimizes ϕ_{ff} is the eigenvector corresponding to the single positive eigenvalue of the generalized eigensystem $\mathbf{A}^t \mathbf{A} \omega = \lambda \mathbf{C} \omega$ [33]. As a final remark notice that the FF method becomes hyperbola specific by replacing matrix \mathbf{C} by $-\mathbf{C}$.

4.2.2 Gradient Weighted Least Squares Fitting

In general the data points used in the conic fitting process are provided by some image processing algorithm such as edge detection. It is reasonable to assume that errors are independent from one point to another, because when detecting a point we usually do not use any information from other points. Moreover it is also reasonable to assume that errors are constant for all points because we use the

4.2. Fitting a Conic Curve to Image Points

same signal processing algorithm. Let the error in a generic point $\mathbf{x}_i = (x_i, y_i)^t$ be Gaussian with zero mean and a 2×2 covariance matrix $\sigma^2 \mathbf{I}$. Notice that the error distribution is assumed to be equal in both directions and uncorrelated. Consider the algebraic distance α_i from point \mathbf{x}_i to the conic ω . The error in the data point affects the measurement of the algebraic distance. Using the result of equation 4.18 comes that the variance of the error in α_i is

$$\sigma_i^2 = \left(\left(\frac{\partial \alpha_i}{\partial x_i} \right)^2 + \left(\frac{\partial \alpha_i}{\partial y_i} \right)^2 \right) \sigma^2 \quad (4.26)$$

An ordinary least square estimator, like the LMS algorithm discussed above, estimates the conic curve ω which minimizes $|\mathbf{A}\omega|$ where \mathbf{A} is the design matrix provided in equation 4.19. It can be shown that the LMS produces the optimal estimation of ω in terms of minimum covariance if equations $\alpha_i = 0$ for $i = 1 \dots M$ have the same variance and are statistically independent [55, 58]. According to equation 4.26 the variance of the error in the algebraic distance α_i depends on the Laplacian of function α on point \mathbf{x}_i . Since the variance on the equations is not constant the LMS estimator does not produce an optimal solution. The estimation result is statistically biased as discussed in [44]. In order to obtain constant variance equations, it is sufficient to divide the algebraic distance function α (equation 4.18) by its gradient. The final normalized objective function becomes

$$\phi_{grad}(\omega) = \sum_{i=1}^M \frac{\alpha_i^2}{\left(\frac{\partial \alpha_i}{\partial x_i} \right)^2 + \left(\frac{\partial \alpha_i}{\partial y_i} \right)^2}$$

The minima of function ϕ_{grad} can not be found by solving an eigensystem as we have done for the LMS, AMS and FF estimators. The objective function is non linear and the problem has not a closed form solution. The minimization process must be performed using iterative gradient descent methods such as Gauss-Newton or Levenberg-Marquardt [58, 27]. As a final remark notice that the AMS estimator described in section 4.2.1 ignores the dependence of the gradient of α on ω in order to obtain a closed form solution for the minimization problem [76, 32].

4. Using Line Images for Calibration Purposes

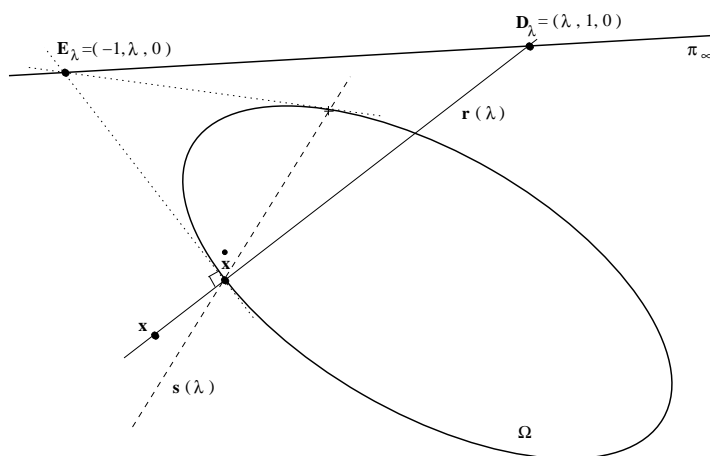


Figure 4.5: The orthogonal distance between a point and a conic

4.2.3 Least Squares Fitting Based on Orthogonal Distances

The big advantage of using algebraic distance for the conic fitting is the gain in computational efficiency. For the methods described in section 4.2.1 the estimation problem can be naturally solved by an eigensystem and a closed form solution is obtained. However there are two major disadvantages in using algebraic distance: a point may contribute differently to the parameter estimation depending on its position on the conics; and the function to minimize is usually not invariant under Euclidean transformations [76]. The approach proposed in section 4.2.2 tries to overcome the former drawback dividing function α by its gradient. Nevertheless the derived objective function is still not invariant to rotation and translation. A natural way to cope with these problems is to replace the algebraic distances by the orthogonal distances which are invariant to Euclidean transformations and do not exhibit the statistical bias described above.

Consider the scheme depicted in Fig. 4.5. The conic curve is described in \mathbb{P}^2 by a 3×3 symmetric matrix Ω . The orthogonal distance β between point x and conic Ω is the smallest Euclidean distance among all distances between x and points in the conic locus. To compute the orthogonal distance β the conic point \hat{x} , closest to x in terms of Euclidean distance, must be determined. Notice that the line joining points x and \hat{x} must be orthogonal to the tangent to the conic Ω on point \hat{x} . Thus, given a generic point x , we start by determining the conic points \hat{x}

4.2. Fitting a Conic Curve to Image Points

for which the stated property is verified. It can be shown that in general there are four different solutions. Given the set of solutions, the Euclidean distance between \mathbf{x} and each candidate point $\dot{\mathbf{x}}$ is computed. The smallest distance corresponds to the orthogonal distance between point \mathbf{x} and conic curve Ω .

The projective plane is euclidean whenever the line at infinity and the circular points are in the canonical position [59, 39, 30]. Consider point $\mathbf{D}_\lambda = (\lambda, 1, 0)^t$ lying on the line at infinity $\pi_\infty = (0, 0, 1)^t$ as depicted in Fig. 4.5. Any two lines intersecting in \mathbf{D}_λ have the same direction and are parallel. Point \mathbf{D}_λ is called the direction point since it represents all possible directions in the plane. Each value of parameter λ corresponds to a different direction. The pencil of lines going through a point \mathbf{x} which is not in π_∞ can be parameterized as follows

$$\begin{aligned} \mathbf{r}(\lambda) &= \mathbf{D}_\lambda \wedge \mathbf{x} \\ &= \tilde{\mathbf{D}}_\lambda \mathbf{x} \end{aligned}$$

$\mathbf{r}(\lambda)$ is the line going through point \mathbf{x} with direction \mathbf{D}_λ . Point $\mathbf{E}_\lambda = (-1, \lambda, 0)^t$ also lies in the line at infinity π_∞ and represents the orthogonal direction to \mathbf{D}_λ . Any two lines intersecting the line at infinity π_∞ on points \mathbf{D}_λ and \mathbf{E}_λ are perpendicular to each other. Consider line $\mathbf{s}(\lambda)$ which is the polar of point \mathbf{E}_λ with respect to conic Ω ($\mathbf{s}(\lambda) = \Omega \mathbf{E}_\lambda$). The intersection point of lines $\mathbf{r}(\lambda)$ and $\mathbf{s}(\lambda)$ is $\mathbf{x}_t(\lambda)$ provided by equation 4.29.

$$\begin{aligned} \mathbf{x}_t(\lambda) &= \mathbf{s}(\lambda) \wedge \mathbf{r}(\lambda) \\ &= (\Omega \tilde{\mathbf{E}}_\lambda) \tilde{\mathbf{D}}_\lambda \mathbf{x} \end{aligned} \tag{4.29}$$

We aim to determine the conic points $\dot{\mathbf{x}}$ such that the corresponding tangent to the conic is perpendicular to the line joining $\dot{\mathbf{x}}$ and \mathbf{x} . Line $\mathbf{s}(\lambda)$ intersects Ω in two points. The tangents to the conic on these points must go through point \mathbf{E}_λ because $\mathbf{s}(\lambda)$ is its polar line. Since \mathbf{E}_λ and \mathbf{D}_λ are orthogonal directions then both tangents are perpendicular to line $\mathbf{r}(\lambda)$. Assume that for a certain value of λ lines $\mathbf{s}(\lambda)$ and $\mathbf{r}(\lambda)$ intersect in a point $\mathbf{x}_t(\lambda)$ lying on the conic curve Ω . According to the stated the line $\mathbf{r}(\lambda)$, joining \mathbf{x} and $\mathbf{x}_t(\lambda)$, is orthogonal to the tangent to the conic on point $\mathbf{x}_t(\lambda)$. Thus, in order to find points $\dot{\mathbf{x}}$, we must solve the following equation in λ

$$\mathbf{x}_t(\lambda)^t \cdot \Omega \cdot \mathbf{x}_t(\lambda) = 0$$

4. Using Line Images for Calibration Purposes

Points $\dot{\mathbf{x}}$ are obtained from equation 4.29 by replacing λ by the roots of the above equation. The equation in λ is of fourth order. This means that there are four λ solutions for which lines $\mathbf{s}(\lambda)$ and $\mathbf{r}(\lambda)$ intersect in a conic point. In other words, there are four conic points $\dot{\mathbf{x}}$ such that the line joining \mathbf{x} and $\dot{\mathbf{x}}$ is perpendicular to the tangent to the conic on $\dot{\mathbf{x}}$. The orthogonal distance β between point \mathbf{x} and conic Ω is the distance from \mathbf{x} to the closest conic point $\dot{\mathbf{x}}$.

The minima ω of the objective function ϕ_{ortho} is the conic curve which fits the data points \mathbf{x}_i by minimizing the sum of the orthogonal distances. Notice that the problem of finding the minima has not a closed form solution. The minimization process must be performed using iterative gradient descent methods such as Gauss-Newton or Levenberg-Marquardt [58, 27]. The conic fitting using orthogonal distances has the disadvantage of requiring a considerable computational effort.

$$\phi_{ortho}(\omega) = \sum_{i=1}^M \beta_i$$

4.2.4 Performance Evaluation

This section compares and evaluates the performance of the conic estimators LMS, AMS, FF, GRAD and ORTHO. The LMS, AMS and FF algorithms are introduced in section 4.2.1. The GRAD method corresponds to the gradient weighted least squares fitting presented in section 4.2.2. The ORTHO algorithm is the conic fitting method based on orthogonal distances (section 4.2.3).

Noise in the Data Points

The performance evaluation is performed using artificial generated data. The test conic is uniformly sampled by 100 points. Two dimensional Gaussian noise with zero mean and standard deviation σ is added to each sample. These samples are the data points used by the different estimators. The principal points of the estimated curve are compared with the ground truth and the mean error is computed over 100 runs of each experiment. Fig 4.6 shows the result of one run for an additive white noise with standard deviation $\sigma = 12$ pixel. The test conic (ground truth) is depicted by the black solid line. The black dots correspond to the data

4.2. Fitting a Conic Curve to Image Points

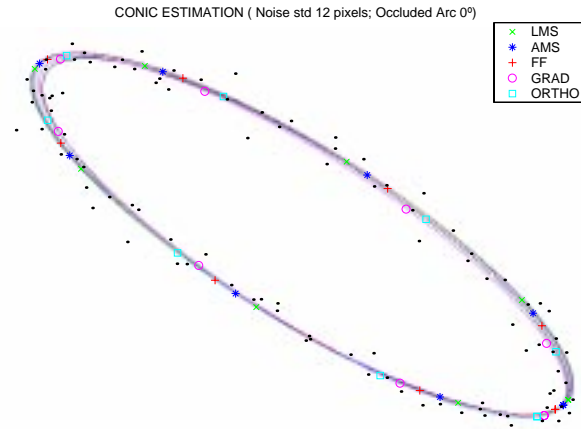


Figure 4.6: Conic fitting in the presence of noise

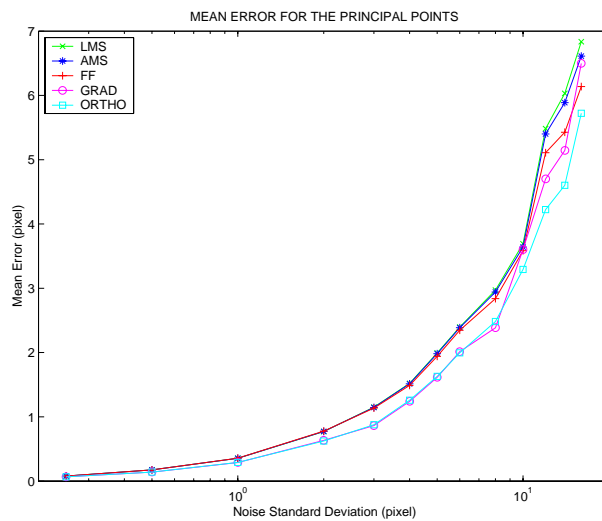


Figure 4.7: The performance of the conic estimators in the presence of noise

4. Using Line Images for Calibration Purposes

points used for the estimation.

Fig. 4.7 shows the performance of the different conic fitting methods in the presence of noise. For all methods the mean estimation error increases when the standard deviation of the noise increases. The performance suffers a graceful degradation in the presence of increasing noise. The GRAD and ORTHO methods are clearly the most robust ones. This is understandable since the LMS, AMS and FF methods are based on the algebraic distance which has some problems as discussed in section 4.2.2. Nevertheless it is important to remind that the GRAD and ORTHO algorithms do not have a closed form solution. The computational effort required to find a solution using gradient descent methods is in general high. Moreover problems of convergence and local minima may always arise when using non linear optimization techniques [27]. Among the methods with closed form solution the FF algorithm seems to be the most robust one. However, since the FF method is ellipse/circle specific, it requires that the type of conic is known in advance.

Partial Occlusion of the Conic

In the previous experiment the entire conic curve is uniformly sampled. The present experiment considers a partial arc of the conic instead of using the entire curve. Consider the problem of fitting a conic to image points. Often the conic curve is not entirely visible in the image plane. Thus it is important to study the behavior of the conic fitting methods when the curve is partially occluded. The data points are generated as follows. An arc of the test conic, with a predefined amplitude, is uniformly sampled by 100 points. The Gaussian noise added to the samples has a standard deviation $\sigma = 2$ pixel. The estimated conic is compared with the ground truth and the mean error is computed over 100 runs for each experiment.

Fig. 4.8 shows the estimation results of the different methods when the sampled arc has an amplitude of 180° , 120° , 90° and 60° . When half of the conic is occluded all the algorithms present a fairly good estimation. For a sampled arc of 120° only the GRAD, ORTHO and AMS methods provide estimations close to the original conic. When the occlusion is greater than 270° none of the methods provide good estimations.

4.2. Fitting a Conic Curve to Image Points

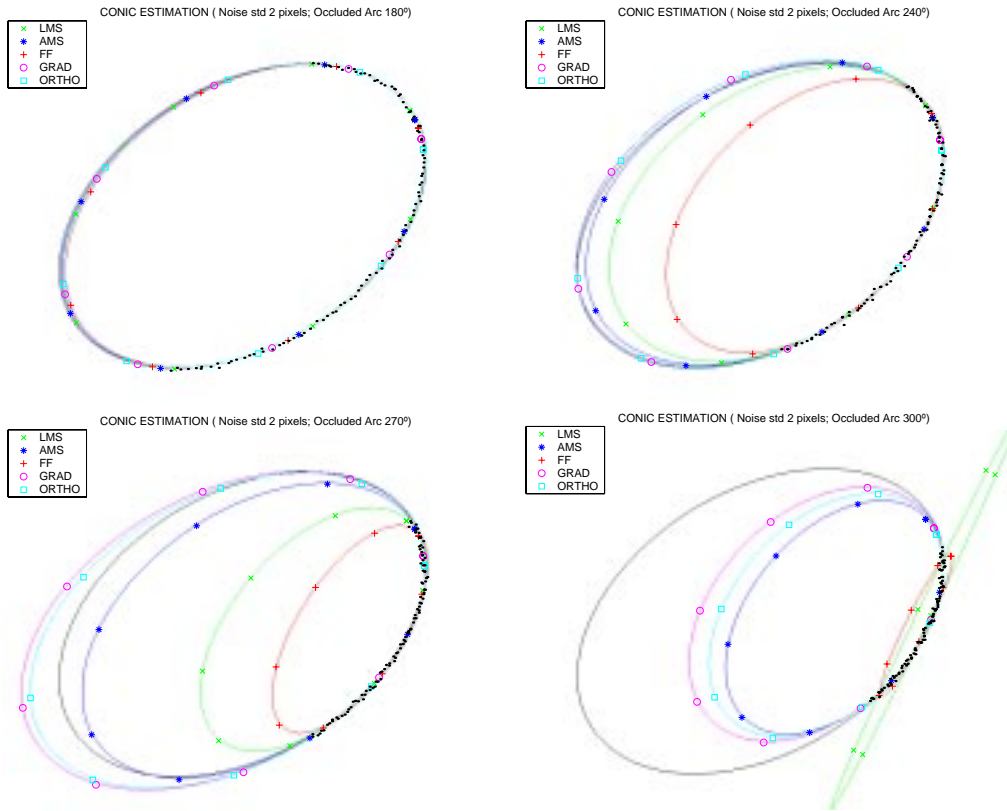


Figure 4.8: Conic fitting in the presence of occlusion

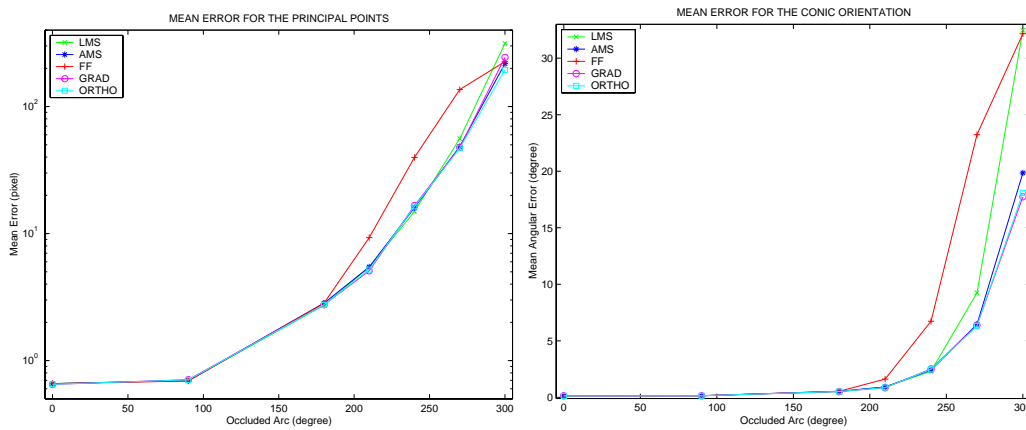


Figure 4.9: The performance of the conic estimators in the presence of occlusion

4. Using Line Images for Calibration Purposes

Fig. 4.9 shows the mean estimation error over 100 runs for each experiment. The graphic on the left refers to the error in the principal points. The graphic on the right shows the error on the orientation angle which is the angle between the major axis of the conic and the plane X axis. As expected an increase on the occlusion angle corresponds to a decrease in the performance of the estimators. For all methods the estimation result suffers an abrupt degradation when the occlusion is higher than 240° . None of them provide an useful estimation when the sampled arc has an amplitude below 100° . Nevertheless the ORTHO, GRAD and AMS methods seem to be the ones that better cope with occlusion. The FF method has clearly the worst performance. The bad behavior of the FF method in the presence of occlusion is in accordance with the experimental observations presented in [33].

4.3 Calibration of Central Catadioptric Systems

In the previous chapter we have proved that an hyperbolic/elliptical system can be calibrated from the image of two lines and that any central catadioptric system can be fully calibrated using a minimum number of three line images. In the former situation the shape of the reflective surface and the relative pose between the camera and the mirror must be known in advance. In the second case there is no requirements besides the system having a single effective viewpoint. The algorithms are summarized in Tab. 3.3 and 3.4.

After estimating the conic curves the calibration procedure is trivial. Section 4.1 presents methods to determine the intersection points between a line and a conic and between a pair of conic curves. The methods are computationally efficient and easy to implement. The crucial step in the calibration procedure is the accurate estimation of the conic curves where the lines are projected. As discussed at the end of chapter 3 this task is hard to accomplish. In general only a small arc of the conic locus is visible in the catadioptric image. According to the results of the previous section the standard conic fitting techniques do not cope well with occlusion.

This section discusses the conditions that must be verified for a set of conic curves to be the catadioptric projection of a set of lines in the scene. The conditions are derived taking into account the geometric properties presented in section 3.3. The idea is to use the properties of central catadioptric line projection to

4.3. Calibration of Central Catadioptric Systems

constrain the search space and improve the performance of the conic fitting in the presence of occlusion.

4.3.1 Calibration of an Hyperbolic/Elliptical System

Consider a pair of lines in the scene imaged by a central catadioptric system combining an hyperbolic/elliptical mirror with a conventional perspective camera. Since the shape of the reflective surface is known then the ξ parameter can be easily determined (Tab. 2.3). Moreover it is assumed that the camera is not rotated with relation to the mirror and that the line at infinity $\hat{\pi}_\infty$ is in the canonical position in the image plane. The system can be calibrated using the strategy presented in section 3.4.1 and summarized in Tab. 3.3. The required geometric construction is schematized on Fig. 3.5 where $\hat{\Omega}_1$ and $\hat{\Omega}_2$ are the conic loci corresponding to the catadioptric line images.

The Necessary Condition $\nu_\xi(\hat{\omega}_1, \hat{\omega}_2) = 0$

Each pair of conic curves $\hat{\Omega}_1, \hat{\Omega}_2$ has a conic $\hat{\Omega}_{12}$ associated with it. The curve $\hat{\Omega}_{12}$, depicted in Fig. 3.5, is a particular realization of the family of line images going through the intersection points \hat{F}_{12} and \hat{B}_{12} . Points \hat{O}, \hat{D}_{12} and \hat{N}_{12} lie on line $\hat{\mu}_{12}$ which is the major axis of $\hat{\Omega}_{12}$. Assume $\check{\xi} = \sqrt{\{\hat{O}\hat{D}_{12}; \hat{N}_{12}\hat{C}_{12}\}}$ with \hat{C}_{12} the conjugate of \hat{M}_{12} with respect to conic $\hat{\Omega}_{12}$ ($\{\hat{M}_{12}, \hat{C}_{12}; \hat{F}_{12}, \hat{B}_{12}\} = -1$). According to the result of proposition 3.5 comes that $\check{\xi}$ is an estimate of parameter ξ which is known. If the pair of conics $\hat{\Omega}_1, \hat{\Omega}_2$ are the catadioptric projection of a pair of lines then $\check{\xi}$ and ξ must be equal.

Consider function ν_ξ provided in equation 4.32. $\hat{\omega}_1, \hat{\omega}_2$ is a pair of conic loci parameterized by points in \mathbf{P}^5 (equation 4.17). The geometric construction of Fig. 3.5 can be performed for any two conics $\hat{\omega}_1, \hat{\omega}_2$. Thus any pair of conic loci has a value $\check{\xi}$ associated with it. Function ν_ξ makes the correspondence between the pairs of conics and $(\xi - \check{\xi})^2$.

$$\nu_\xi(\hat{\omega}_1, \hat{\omega}_2) = (\xi - \check{\xi})^2 \quad (4.32)$$

The domain of function ν_ξ is the space of pairs of conic curves. If $\hat{\omega}_1, \hat{\omega}_2$ is a pair of catadioptric line images then $\nu_\xi(\hat{\omega}_1, \hat{\omega}_2)$ must be null. Since ν_ξ is always

4. Using Line Images for Calibration Purposes

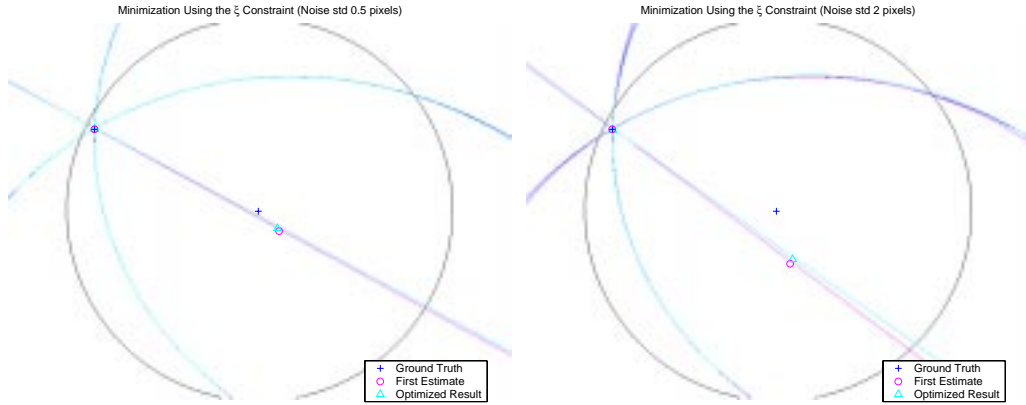


Figure 4.10: Calibration of an hyperbolic sensor from 2 line images

positive, the pairs of catadioptric line projections are global minima of the function. Notice however that condition $\nu_{\xi}(\hat{\omega}_1, \hat{\omega}_2) = 0$ is necessary but not sufficient for $\hat{\omega}_1, \hat{\omega}_2$ to be a pair of catadioptric line images. This means that function ν_{ξ} may have other minima than the ones corresponding to the line projections.

The Calibration Procedure

Fig. 4.10 exhibits two simulated test images of a pair of lines acquired by an hyperbolic sensor. The blue cross signs the image center and the black circle is the image of the mirror boundary. The lines in the scene are mapped into conic curves which are marked using dark blue (the ground truth). The data points used to estimate the line images are taken from the conic arcs inside the circle. Two dimensional zero mean gaussian noise is added to the image points. On the left the noise standard deviation is $\sigma = 0.5$ pixels, and on the right the standard deviation is $\sigma = 2$ pixels.

An initial estimate of $\hat{\omega}_1$ and $\hat{\omega}_2$ can be obtained by fitting a conic curve to the data points visible in the image plane. The conic fitting is performed using the AMS method described in section 4.2.1. The AMS method is suitable for the task since it has a closed form solution and it is more robust to occlusion than the LMS and FF algorithms. The initial estimates of the conic curves are marked by the magenta line. Notice that inside the circle the arcs of the estimated conics are coincident with the ground truth.

4.3. Calibration of Central Catadioptric Systems

The initial estimates of conics $\hat{\omega}_1$ and $\hat{\omega}_2$ are used to calibrate the catadioptric system using the procedure summarized on Tab. 3.3. The computed image center is signed by the small magenta circle in both images of Fig. 4.10. The calibration results are incorrect in both cases. This is due to a wrong estimation of the catadioptric line images.

Consider function ϵ_h provided in equation 4.33. Function ϕ is the sum of the square of the algebraic distances between the curve and the corresponding data points (equation 4.20). According to the discussion above, if $\hat{\omega}_1, \hat{\omega}_2$ correspond to a pair of catadioptric line images then $\check{\xi}$ and ξ are the same and $\nu_\xi(\hat{\omega}_1, \hat{\omega}_2)$ must be null. The difference $(\xi - \check{\xi})^2$ is introduced as a penalty term weighted by a multiplying parameter k_0 . The conic loci can be estimated by minimizing function ϵ_h . The function is highly non linear and the minima must be found using iterative gradient descending methods [58, 27].

$$\epsilon_h(\hat{\omega}_1, \hat{\omega}_2) = \phi(\hat{\omega}_1) + \phi(\hat{\omega}_2) + k_0 \nu_\xi(\hat{\omega}_1, \hat{\omega}_2) \quad (4.33)$$

The entire procedure can be summarized in the following manner. An initial estimate of the catadioptric line images is obtained from the data points by applying the AMS algorithm. The image center \hat{O} and the image of the absolute conic $\hat{\Omega}_\infty$ are determined following the steps of Tab. 3.3. The value of $\check{\xi}$ is computed using the cross-ratio relation of proposition 3.5. A novel estimation of the line images is obtained by minimizing function ϵ_h where $(\xi - \check{\xi})^2$ is introduced as a penalty term. The procedure is iterated until the value of ϵ_h is below a certain threshold.

The conic loci estimated using this procedure are depicted in Fig. 4.10. The cyan triangle symbolizes the image center computed following the steps of Tab. 3.3. In both images the new estimate of the image center is closer to the ground truth than the initial one. Nevertheless the improvement is slight and the calibration results are far from being correct.

Conclusions

The calibration procedure performs the minimization of function ϵ_h using gradient descendent methods. If the objective function has local minima and/or saddle points then there is no way to guarantee that the iterative process converges to the

4. Using Line Images for Calibration Purposes

right solution [58, 27]. In this case the gradient descending method may stop on one of these points instead of converging to the ground truth. This may explain the poor calibration results observed in Fig. 4.10.

The domain of function ϵ_h is the space of pairs of conic curves. Since the domain has dimension greater than 3 the function can not be plotted in a straightforward manner. Due to the occlusion problem there are several pairs of conic curves $\hat{\omega}_1, \hat{\omega}_2$ which minimize the algebraic distance to the data points. This is the reason why the AMS method is unable to provide a good initial estimate of the conic loci. If it is true that any pair of lines projects into a pair of conics, it is not true that any pair of conic loci is the catadioptric image of a pair of lines. The idea of using function ν_ξ as a penalty term is to avoid solutions corresponding to conic loci that can not be line images. However the condition $\nu_\xi(\hat{\omega}_1, \hat{\omega}_2) = 0$ is necessary but not sufficient. Not all minima of ν_ξ are pairs of catadioptric line images. Thus it is reasonable to assume that the objective function ϵ_h (equation 4.32) has local minima and/or saddle points where the iterative gradient descending method stops.

4.3.2 Calibration of a General Central Catadioptric System

Any central catadioptric system can be fully calibrated from the image of three lines. The calibration method is summarized on Tab. 3.4 and the required geometric construction is depicted in Fig. 3.6 and 3.7.

Fig. 4.11 shows a simulated image of three lines in the scene. The lines are imaged into three conic curves marked using dark blue. Each conic arc visible in the image is sampled by 300 points. Two dimensional zero mean gaussian noise with standard deviation σ is added to the data points. In the images of the top row the noise standard deviation is $\sigma = 0.5$, and in the bottom row is $\sigma = 2$.

The magenta conic loci are the initial estimations of the catadioptric line images (right images in Fig. 4.11). These estimates are obtained by fitting conics to the data points using the AMS method (section 4.2.1). According to the result of proposition 3.8 the image center \hat{O} must be collinear with the intersection points of any two line images. As you can observe this does not happen which proves that the conic loci are not correctly estimated.

The AMS algorithm fits the conic to the data points by minimizing the sum of

4.3. Calibration of Central Catadioptric Systems

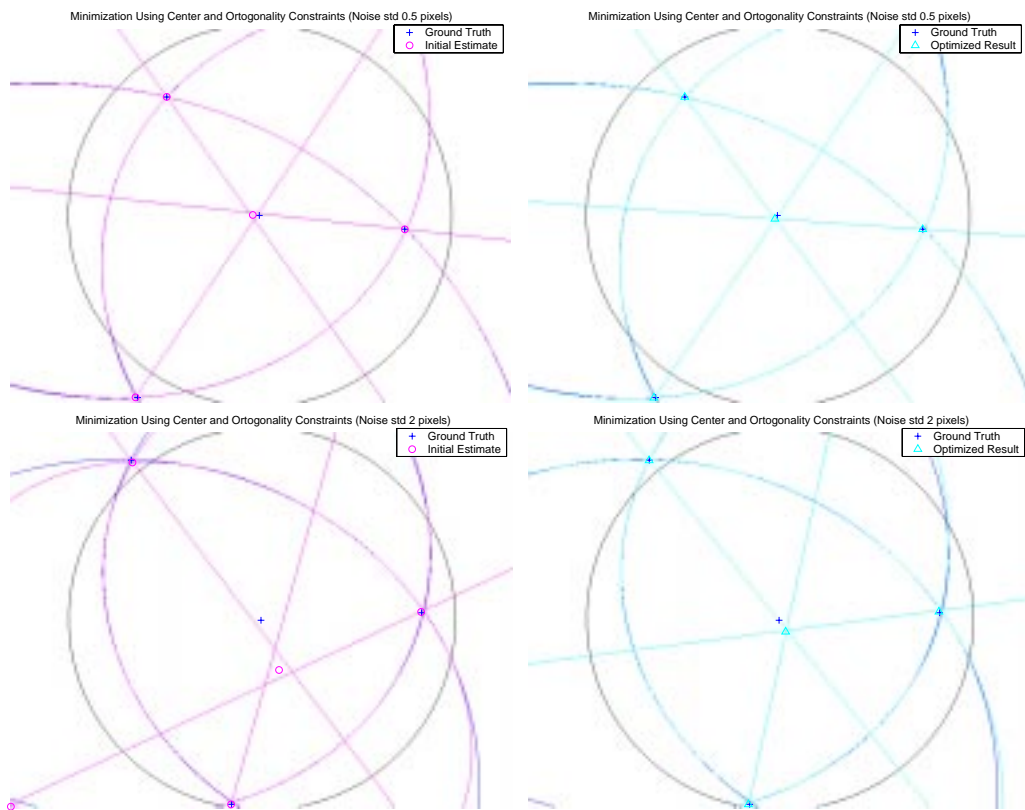


Figure 4.11: Calibration of a general catadioptric sensor from 3 line images

4. Using Line Images for Calibration Purposes

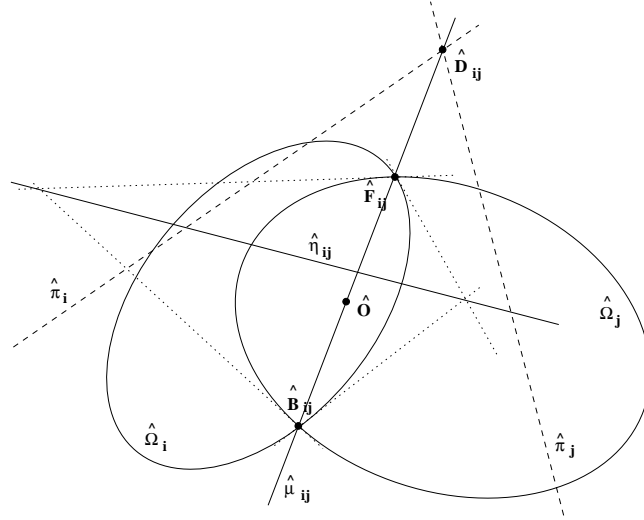


Figure 4.12: Geometric properties of a pair of catadioptric line images

the squares of the algebraic distances. Since the curve is partially occluded there are several minima in the objective function. We aim to overcome the problem by deriving conditions that must be verified by a set of catadioptric line images. The idea is to include these conditions as penalty terms and obtain an objective function with a single minimum in the correct solution.

The Image Center Constraint

Assume that conics $\hat{\Omega}_i, \hat{\Omega}_j$ in Fig. 4.12 are the catadioptric images of a pair of lines. The conic curves intersect each other on points $\hat{F}_{ij}, \hat{B}_{ij}$. According to the result of proposition 3.8 the image center \hat{O} must lie on line $\hat{\mu}_{ij}$ going through the intersection points $\hat{F}_{ij}, \hat{B}_{ij}$. Assume that $\hat{\pi}_i, \hat{\pi}_j$ are the polar lines of the image center \hat{O} with respect to conics $\hat{\Omega}_i, \hat{\Omega}_j$. The polar lines intersect on point \hat{D}_{ij} . From corollary 2.4 comes that \hat{D}_{ij} must lie on line $\hat{\nu}_{ij}$. Thus if $\hat{\Omega}_i, \hat{\Omega}_j$ are the catadioptric projections of a pair of lines then points $\hat{O}, \hat{D}_{ij}, \hat{F}_{ij}$ and \hat{B}_{ij} must be collinear.

Consider the pair of conic loci $\hat{\Omega}_i, \hat{\Omega}_j$ and point \hat{O} . Since the polar lines are $\hat{\pi}_i = \hat{\Omega}_i \hat{O}$ and $\hat{\pi}_j = \hat{\Omega}_j \hat{O}$ then the intersect point is $\hat{D}_{ij} = \hat{\pi}_i \wedge \hat{\pi}_j = (\hat{\Omega}_i \hat{O}) \wedge (\hat{\Omega}_j \hat{O})$. Notice that \hat{D}_{ij} and \hat{O} are conjugate with respect to both conics. Consider line $\hat{\mu}_{ij}$ defined by points \hat{O} and \hat{D}_{ij} . Any point P lying on $\hat{\mu}_{ij}$ can be written as a

4.3. Calibration of Central Catadioptric Systems

linear combination of points $\hat{\mathbf{O}}$ and $\hat{\mathbf{D}}_{ij}$ (equation 4.34).

$$\mathbf{P}(\lambda) = \hat{\mathbf{O}} + \lambda \hat{\mathbf{D}}_{ij} \quad (4.34)$$

Line $\hat{\mu}_{ij}$ intersects $\hat{\Omega}_i$ in two points. The intersection points can be determined by solving equation $\mathbf{P}(\lambda)^t \hat{\Omega}_i \mathbf{P}(\lambda) = 0$ in order to λ . Replacing \mathbf{P} by the result of equation 4.34 and taking into account that $\hat{\mathbf{D}}_{ij}$ and $\hat{\mathbf{O}}$ are conjugate with respect to $\hat{\Omega}_i$ yields

$$\lambda_i = \pm \sqrt{-\frac{\hat{\mathbf{O}}^t \hat{\Omega}_i \hat{\mathbf{O}}}{\hat{\mathbf{D}}_{ij}^t \hat{\Omega}_i \hat{\mathbf{D}}_{ij}}}$$

The same procedure can be used to determine the points where line $\hat{\mu}_{ij}$ and conic $\hat{\Omega}_j$ intersect. In this case the λ solution is

$$\lambda_j = \pm \sqrt{-\frac{\hat{\mathbf{O}}^t \hat{\Omega}_j \hat{\mathbf{O}}}{\hat{\mathbf{D}}_{ij}^t \hat{\Omega}_j \hat{\mathbf{D}}_{ij}}}$$

Assume that $\hat{\Omega}_i, \hat{\Omega}_j$ are the catadioptric projections of a pair of lines and $\hat{\mathbf{O}}$ is the image center. From the discussion above comes that line $\hat{\mu}_{ij}$ intersects both conics in the same pair of points. These points are the locus where the conics $\hat{\Omega}_i, \hat{\Omega}_j$ intersect each other (points $\hat{\mathbf{F}}_{ij}, \hat{\mathbf{B}}_{ij}$). Since the pairs of points $\mathbf{P}(\lambda_i^+), \mathbf{P}(\lambda_j^+)$ and $\mathbf{P}(\lambda_i^-), \mathbf{P}(\lambda_j^-)$ are coincident then the solutions λ_i and λ_j must be equal. Thus,

$$\lambda_i^2 = \lambda_j^2 \Leftrightarrow \hat{\mathbf{D}}_{ij}^t (\hat{\Omega}_j \hat{\mathbf{D}}_{ij} \hat{\mathbf{O}}^t \hat{\Omega}_i - \hat{\Omega}_i \hat{\mathbf{D}}_{ij} \hat{\mathbf{O}}^t \hat{\Omega}_j) \hat{\mathbf{O}} = 0$$

Assume function χ provided in equation 4.38 with $\hat{\omega}_i, \hat{\omega}_j$ a pair of conic curves parameterized in \mathbf{P}^5 . If $\hat{\omega}_i, \hat{\omega}_j$ correspond to the catadioptric projection of a pair of lines and $\hat{\mathbf{O}}$ is the image center then $\chi(\hat{\omega}_i, \hat{\omega}_j)$ must be null.

$$\chi(\hat{\omega}_i, \hat{\omega}_j) = (\hat{\mathbf{D}}_{ij}^t (\hat{\Omega}_j \hat{\mathbf{D}}_{ij} \hat{\mathbf{O}}^t \hat{\Omega}_i - \hat{\Omega}_i \hat{\mathbf{D}}_{ij} \hat{\mathbf{O}}^t \hat{\Omega}_j) \hat{\mathbf{O}})^2 \quad (4.38)$$

4. Using Line Images for Calibration Purposes

Orthogonality Constraints

Consider the pair of conic curves $\hat{\Omega}_i, \hat{\Omega}_j$ depicted in Fig. 4.12. $\hat{\eta}_{ij}$ is the line defined by the poles of $\hat{\mu}_{ij}$ with respect to conics $\hat{\Omega}_i, \hat{\Omega}_j$. According to proposition 3.9, if the conic curves are two catadioptric line images then $\hat{\eta}_{ij}$ is the locus where the corresponding line of normals is mapped. This line is defined by the normal directions to the planes containing the imaged lines and the system effective viewpoint (Fig. 3.4). Consider the intersection point \hat{D}_{ij} and the image of the absolute conic $\hat{\Omega}_\infty$. Since \hat{D}_{ij} is the locus of the common direction of the planes containing the imaged lines (corollary 2.4) then the polar line of \hat{D}_{ij} with respect to $\hat{\Omega}_\infty$ is $\hat{\eta}_{ij}$.

Assume function ν defined on equation 4.39 where $\hat{\Omega}_\infty^*$ is the conic envelope of the image of the absolute conic. If conics $\hat{\Omega}_i, \hat{\Omega}_j$ are the catadioptric projection of a pair of lines then the pole of $\hat{\eta}_{ij}$ with respect to the image of the absolute conic is \hat{D}_{ij} . The vector product $\hat{D}_{ij} \wedge (\hat{\Omega}_\infty^* \hat{\eta}_{ij})$ is zero and function $\nu(\hat{\omega}_i, \hat{\omega}_j)$ must be null.

$$\nu(\hat{\omega}_i, \hat{\omega}_j) = (\hat{D}_{ij} \wedge (\hat{\Omega}_\infty^* \hat{\eta}_{ij}))^t (\hat{D}_{ij} \wedge (\hat{\Omega}_\infty^* \hat{\eta}_{ij})) \quad (4.39)$$

Calibration Procedure

Consider the catadioptric image of three lines depicted in Fig. 4.11. The lines are projected on conic curves $\hat{\Omega}_1, \hat{\Omega}_2$ and $\hat{\Omega}_3$ which are represented in \mathbf{P}^5 by points $\hat{\omega}_1, \hat{\omega}_2$ and $\hat{\omega}_3$. The initial estimates of the conic loci are obtained using the AMS algorithm. The image center \hat{O} and the image of the absolute conic $\hat{\Omega}_\infty$ are estimated following the steps summarized on Tab. 3.4. In general the initial conic estimates are not accurate and lines $\hat{\mu}_{12}, \hat{\mu}_{13}$ and $\hat{\mu}_{23}$ do not intersect in the same point (Fig. 3.6). In this case the image center is determined using normal least squares [58]. A novel estimation for the catadioptric line images is obtained by minimizing function ϵ_g provided in equation 4.40. The image center and the absolute conic are determined using the new conic estimations. The procedure is iterated until function ϵ_g takes values below a certain threshold.

4.3. Calibration of Central Catadioptric Systems

$$\epsilon_g(\hat{\omega}_1, \hat{\omega}_2, \hat{\omega}_3) = \sum_{i=1}^3 \phi(\hat{\omega}_i) + k_0 \sum_{ij} \chi(\hat{\omega}_i, \hat{\omega}_j) + k_1 \sum_{ij} \nu(\hat{\omega}_i, \hat{\omega}_j) \quad (4.40)$$

The estimated conics $\hat{\Omega}_1$, $\hat{\Omega}_2$ and $\hat{\Omega}_3$ are used to calibrate the system (Tab. 3.4). The final results can be observed on the right of Fig. 4.11. The lines defined by the conic intersections meet in a single point for both noise level situations ($\sigma = 0.5$ and $\sigma = 2$). As you can observe the intersection point is close to the real image center.

Conclusions

The domain of functions χ and ν , provided in equations 4.38 and 4.39, is the space of pairs of conics. We have shown that if $\hat{\omega}_i$, $\hat{\omega}_j$ are the catadioptric images of a pair of lines then both $\chi(\hat{\omega}_i, \hat{\omega}_j)$ and $\nu(\hat{\omega}_i, \hat{\omega}_j)$ must be null. Remark that conditions $\chi(\hat{\omega}_i, \hat{\omega}_j) = 0$ and $\nu(\hat{\omega}_i, \hat{\omega}_j) = 0$ are necessary but not sufficient. Both functions may have minima which do not correspond to pairs of catadioptric line images.

Function ϵ_g (equation 4.40) is highly non linear and the minima must be found using iterative gradient descending methods [58, 27]. As stated in the previous section convergence problems may arise whenever the objective function has local minima and/or saddle points. Functions χ and ν are used as penalty terms in function ϵ_g . The goal is to avoid solutions, which minimize the algebraic distances to the data points, but do not correspond to a coherent catadioptric image of a set of lines in the scene.

We have performed several experiments like the one depicted in Fig. 4.11. The proposed approach worked properly in many of the simulations. The conic loci were accurately estimated and the system was correctly calibrated. Nevertheless we have also detected many situations for which the procedure did not work well. Notice that we have no guarantee that the objective function ϵ_g has not multiple minima and/or saddle points. As stated in the previous section the gradient descending method may converge to one of these points which do not correspond to the correct conic loci where the lines are imaged.

4. Using Line Images for Calibration Purposes



Figure 4.13: Segmentation of catadioptric line images

4.3.3 Experiments with Real Images

Experiments with real images are reported in this section. Fig. 4.13 depicts the graphic user interface (GUI) used to segment the line images and obtain the data points to perform the conic fitting. An edge detector is applied to the original catadioptric image [21]. The conic curves corresponding to line images are segmented by hand.

Fig. 4.14 is an example of a test image. The GUI is used to obtain the data points corresponding to 5 line images. The initial estimation of the conic loci is obtained using the AMS method. The result is depicted on the left image where the estimated conics perfectly fit the visible arcs of the catadioptric line images. Nevertheless proposition 3.8 is not verified which means that the conic curves are not correctly estimated. The right image shows the final estimation of the conic curves obtained by minimizing function ϵ_g (equation 4.40) following the procedure of section 4.3.2. This set of catadioptric line images seems coherent since it verifies the geometric properties derived in chapter 3. The system is calibrated in a straightforward manner following the steps summarized in Tab. 3.4.

Fig. 4.14 and 4.15 show the same test image. In the former 5 lines are used to calibrate the catadioptric systems. In Fig. 4.15 we are using an additional line for the same purpose. The left image shows the initial estimation of the 6 catadioptric

4.3. Calibration of Central Catadioptric Systems

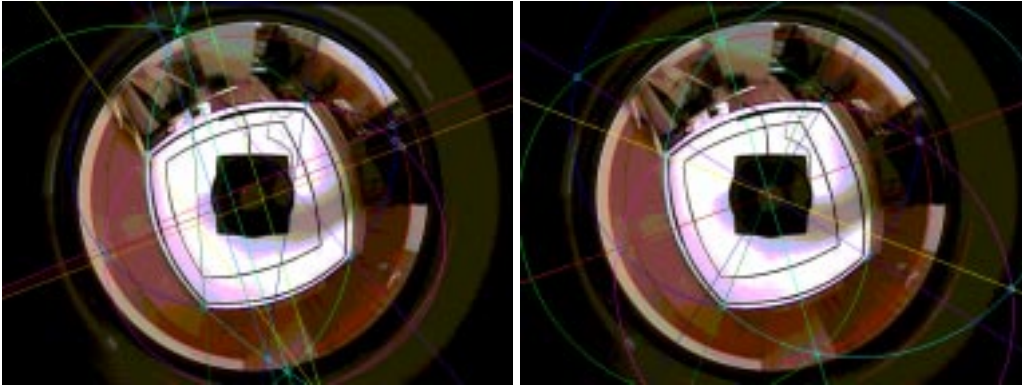


Figure 4.14: Calibration of a general catadioptric system

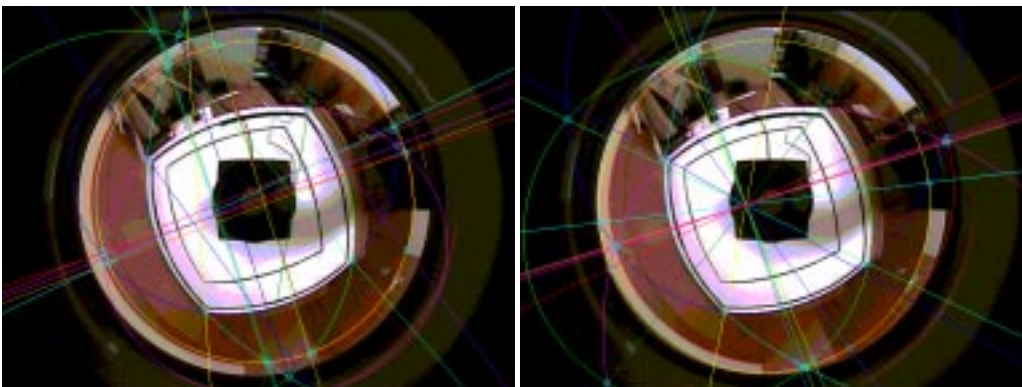


Figure 4.15: Convergence for a wrong minima

4. Using Line Images for Calibration Purposes

line images obtained using the AMS method. The right image exhibits the final estimation of the conic loci determined by minimizing function ϵ_g (section 4.3.2). The lines defined by the conic intersections meet in a single point as stated in proposition 3.8. Compare the left image of Fig. 4.15 with the right image of Fig. 4.15. Notice that same line images have distinct estimations. Moreover different calibration results are obtained for the same catadioptric systems using the same test image. This can only be explained by convergence problems during the iterative minimization due to the existence of multiple minima.

4.4 Closure

Chapter 3 proves that in general a line in the scene is mapped in a conic locus in the catadioptric image plane. Moreover it shows that an hyperbolic/elliptical system can be calibrated from the image of two lines and that any central catadioptric sensor can be fully calibrated using a minimum of three line images.

The present chapter starts by reviewing issues on the geometry of conic curves. Algorithms to compute the intersection between a line and a conic and between a pair of conics are proposed. It is shown that estimating catadioptric line images from image points is far from being a trivial task. In general only small arcs of the line images are visible in the catadioptric image plane. Since standard conic fitting techniques do not perform well in the presence of strong occlusion, the accurate estimation of the conic loci is hard to accomplish.

The calibration algorithms of tables 3.3 and 3.4 require the knowledge of the conic loci where lines are mapped. Due to the occlusion problem these conic curves can not be accurately estimated using standard conic fitting techniques. Section 4.3 derives geometric conditions that must be verified by a set of conic loci to be the catadioptric image of a set of lines. We try to use these necessary conditions to constrain the search space and improve the estimation accuracy. However, since the resulting objective function must be minimized using gradient descending techniques, convergence problems may arise. Simulation results show that the calibration of an hyperbolic/elliptical sensor using just two line images is in general poor. The objective function presents local minima and/or saddle points where the iterative minimization process stops. A similar problem is observed when using three or more lines to calibrate. In some of the experiments the sen-

4.4. Closure

sor is successfully calibrated. In others the set of catadioptric line images is not correctly estimated.

4. Using Line Images for Calibration Purposes

Chapter 5

The Paracatadioptric Camera

The paracatadioptric sensor combines a parabolic shaped mirror and a camera inducing an orthographic projection. Such a configuration provides a wide field of view while keeping a single effective viewpoint. As discussed in chapter 2, there are other catadioptric configurations providing central panoramic imaging [5]. Panoramic central catadioptric systems can be built by combining an hyperbolic mirror with a perspective camera and, a parabolic mirror with an orthographic camera. The construction of the former requires a careful alignment between the mirror and the imaging device. The camera projection center must be positioned in the outer focus of the hyperbolic reflective surface. The paracatadioptric camera is easier to construct being broadly used in applications requiring omnidirectional vision [67, 64, 20, 56].

The present chapter focuses exclusively on paracatadioptric sensors. A general mapping model for central catadioptric image formation has been introduced in chapter 2. The central catadioptric projection is isomorphic to a projective mapping from a sphere, centered in the effective viewpoint, to a plane with projection center on the perpendicular to the plane [34, 7]. For the particular case of paracatadioptric sensors the projection center lies on the sphere and the projective mapping is a stereographic projection. The plane and the final catadioptric image are related by a affine transformation depending on the mirror and camera intrinsic parameters. Section 5.1 reviews the paracatdioptric image formation model. Due to its particular features the paracatadioptric sensor has properties which are not verified by other types of central catadioptric systems (section 3.3.2).

5. The Paracatadioptric Camera

Section 5.2 introduces an effective way to calibrate the paracatadioptric camera using lines. It has already been proved that three line images are enough to calibrate any central catadioptric system (section 3.4.2) [8]. However, as discussed in chapter 4, the estimation of the conic curves where lines are mapped is hard to accomplish. In section 5.2 we show that a set of conic curves corresponds to paracatadioptric line images if, and only if, certain properties are verified. These properties are used to constrain the search space and correctly estimate the curves. If the camera is skewless and the aspect ratio is known then the conic fitting problem is solved naturally by an eigensystem. For the general situation the conic curves are estimated using non-linear optimization.

Section 5.3 proposes a conic fitting method to estimate lines in the paracatadioptric image plane. The algorithm is specific for line images and requires that the calibration of the parabolic system is known. If it is true that any line is projected into a conic, it is not true that any conic is the image of a line (section 3.3.2). Considering the space of all conic curves, the paracatadioptric line images lie on a linear subspace which depends on the system calibration. We show that the line images can be accurately determined by constraining the search space. The corresponding locus is estimated by fitting a conic in the subspace to the data points. The approach is computationally efficient since the fitting problem can be solved by an eigensystem.

5.1 Paracatadioptric Camera Model

This section reviews the image formation model for the paracatadioptric system. The model presented in here is the particular case of the general mapping model for central catadioptric systems when the ξ parameter is unitary (see Tab. 2.3).

Assume a paracatadioptric system combining a parabolic mirror with latus rectum $4p$, and an orthographic camera. The principal axis of the camera must be aligned with the symmetry axis of the paraboloid. The paracatadioptric projection can be modeled by a stereographic projection from an unitary sphere, centered in the effective viewpoint, into a plane Π_∞ as shown in Fig. 5.1.

The world point shown in Fig. 5.1 is imaged at point \hat{x} in the paracatadioptric image plane. The mapping can be described as follows. To each visible scene point corresponds an oriented projective ray $\mathbf{x} = (x, y, z)^t$, joining the 3D point

5.1. Paracatadioptric Camera Model

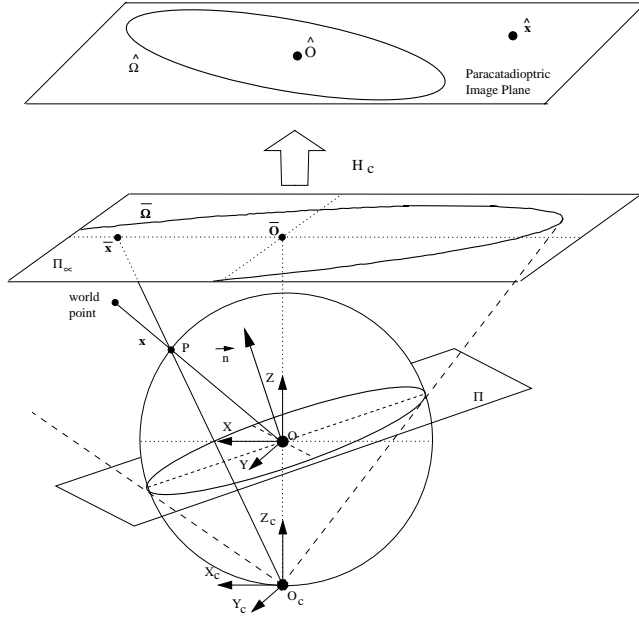


Figure 5.1: Model for paracatadioptric image formation

with the projection center O . The projective ray intersects the unit sphere in a single point X_c . Consider a point O_c , with coordinates $(0, 0, -1)^t$, which lies on the unitary sphere. To each x corresponds an oriented projective ray \bar{x} joining O_c with the intersection point X_c . The non-linear mapping \bar{h} (equation 5.1) corresponds to projecting the scene in the unity sphere surface and then re-projecting the points on the sphere into a plane Π_∞ from the novel projection center O_c . Points in catadioptric image plane \hat{x} are obtained after a collineation H_c of 2D projective points \bar{x} . Equation 5.2 shows that the affine transformation H_c depends on the intrinsic parameters K_c of the orthographic camera, and on the latus rectum of the parabolic mirror.

$$\bar{h}(x) = (x, y, z + \sqrt{x^2 + y^2 + z^2})^t \quad (5.1)$$

$$\hat{x} = \underbrace{K_c \begin{bmatrix} 2p & 0 & 0 \\ 0 & 2p & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{H_c} \bar{x} \quad (5.2)$$

5. The Paracatadioptric Camera

Consider the plane $\Pi = (\mathbf{n}, 0)^t$ going through the effective viewpoint \mathbf{O} as depicted in Fig. 5.1 ($\mathbf{n} = (n_x, n_y, n_z)^t$). The paracatadioptric image of any line lying on Π is the conic curve $\hat{\Omega}$. The line in the scene is projected into a great circle in the sphere surface. This great circle is the curve of intersection of plane Π , containing both the line and the projection center \mathbf{O} , and the unit sphere. The projective rays $\bar{\mathbf{x}}$, joining \mathbf{O}_c to points in the great circle, form a central cone surface. The central cone, with vertex in \mathbf{O}_c , projects into the conic $\bar{\Omega}$ in plane Π_∞ (equation 5.3). Since the image plane and Π_∞ are related by collineation \mathbf{H}_c , the result of equation 5.4 comes in a straightforward manner.

$$\bar{\Omega} = \begin{bmatrix} -n_z^2 & 0 & n_x n_z \\ 0 & -n_z^2 & n_y n_z \\ n_x n_z & n_y n_z & n_z^2 \end{bmatrix} \quad (5.3)$$

$$\hat{\Omega} = \begin{bmatrix} a & b & d \\ b & c & e \\ d & e & f \end{bmatrix} = \mathbf{H}_c^{-t} \bar{\Omega} \mathbf{H}_c^{-1} \quad (5.4)$$

5.2 Paracatadioptric Camera Calibration Using Lines

It has already been proved that any central panoramic system can be fully calibrated from the image of three lines in general position [8]. However, since lines are mapped into conic curves which are only partially visible, the accurate estimation of catadioptric line images is far from being a trivial task

The present section focuses on paracatadioptric camera calibration using lines in general position. If it is true that any line maps into a conic in the catadioptric image plane, it is not true that any conic is the image of a line. We derive for the first time the necessary and sufficient conditions that must be verified by a set of conic curves to be the paracatadioptric projection of lines. We also show that the derived conditions can be used to accurately estimate the line images by non-linear optimization. Moreover if the system is skewless and the aspect ratio is known then the lines can be computed by solving an eigensystem. Given the image of at least three lines the paracatadioptric camera is easily calibrated using the algorithm presented in Tab. 3.4.

Other authors have already proposed algorithms to calibrate a paracatadioptric

5.2. Paracatadioptric Camera Calibration Using Lines

camera [35, 45, 74]. The approach presented in [45] requires a sequence of paracatadioptric images. The system is calibrated using the consistency of pairwise tracked point features across the sequence, based on the characteristics of catadioptric imaging. In [74], the center and focal length are determined by fitting a circle to the image of the mirror boundary. The method is simple and can be easily automated, however it is not very accurate and requires the visibility of the mirror boundary. Its major drawback is that it is only applicable for the situation of a skewless camera with unitary aspect ratio. Geyer and Daniilidis propose a calibration algorithm using line images [35]. They present a closed-form solution for focal length, image center, and aspect ratio for skewless cameras, and a polynomial root solution in the presence of skew. The line images are estimated taking into account the properties of parabolic projections. Nevertheless the conic curves verifying those properties are not necessarily the paracatadioptric projection of lines. We will return to this discussion and compare the approach of [35] with the one proposed in here.

5.2.1 Calibration Algorithm

The paracatadioptric system is calibrated whenever the collineation \mathbf{H}_c is known. The algorithm presented on chapter 3 and summarized in Tab. 3.4 can be used to calibrate any central catadioptric system from a minimum of three line images. If the sensor is parabolic then the ξ parameter is unitary and the optical axis of the camera must be aligned with the symmetry axis of the mirror. The camera can not be rotated with relation to the mirror ($\mathbf{R}_c = \mathbf{I}$ in equation 2.7) and the transformation \mathbf{H}_c must be affine (equation 5.2). Assume that the image center is $\mathbf{C} = (c_x, c_y)^t$, and that α^2 , f_o and s_k are the aspect ratio, the focal length and the skew of the orthographic imaging device. The affine transformation \mathbf{H}_c is provided in equation 5.5 where $f_c = 2f_o p$ is a measurement in pixels of the combined focal length of the camera and the mirror.

$$\mathbf{H}_c = \begin{bmatrix} \alpha f_c & s_k & c_x \\ 0 & \alpha^{-1} f_c & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (5.5)$$

Tab. 5.1 summarizes the steps to calibrate a paracatadioptric system from

5. The Paracatadioptric Camera

Step 1	Determine the catadioptric line images $\hat{\Omega}_i$ for $i = 1, 2, 3 \dots K$
Step 2	For each pair of conics $\hat{\Omega}_i, \hat{\Omega}_j$, compute the intersection points $\hat{F}_{ij}, \hat{B}_{ij}$ and determine the corresponding line $\hat{\mu}_{ij} = \hat{F}_{ij} \wedge \hat{B}_{ij}$
Step 3	Estimate the image center \hat{O} which is the intersection point of lines $\hat{\mu}_{ij}$.
Step 4	For each conic $\hat{\Omega}_i$ compute the polar line $\hat{\pi}_i$ of the image center \hat{O} ($i = 1, 2, 3 \dots K$).
Step 5	For each conic curve obtain the points \hat{I}_i and \hat{J}_i where line $\hat{\pi}_i$ intersects $\hat{\Omega}_i$ ($i = 1, 2, 3 \dots K$)
Step 6	Estimate the conic $\hat{\Omega}_\infty$ going through points \hat{I}_i, \hat{J}_i ($i = 1, 2, 3 \dots K$)
Step 7	Perform the Cholesky decomposition of $\hat{\Omega}_\infty$ to estimate matrix \mathbf{H}_c

Table 5.1: Calibrating a paracatadioptric system using K lines ($K \geq 3$)

the image of K lines in general position. The algorithm is an adaptation of the general method discussed in section 3.4.2. The six last steps in Tab. 3.4 have been omitted since the ξ parameter is already known to be unitary. Notice that \mathbf{H}_c can be determined from the Cholesky decomposition of $\hat{\Omega}_\infty$ because \mathbf{H}_c is an upper triangular matrix (equation 5.5) and $\hat{\Omega}_\infty = \mathbf{H}_c^{-t} \mathbf{H}_c^{-1}$.

The calibration of the paracatadioptric system is straightforward whenever the conic curves corresponding to the line images are known. However the estimation of these conics using image points is hard to accomplish as discussed in chapter 4. There are several algorithms to fit a conic curve to data points. A robust conic fitting algorithm has to cope with noisy data points, biasing due to curvature and partial occlusions. The occlusion problem is of particular importance for our purposes. By occlusion we mean that the available data points lie on a small arc of the curve. It is intuitive that in these circumstances, even for small amounts of

5.2. Paracatadioptric Camera Calibration Using Lines

noise, it is very hard to obtain the correct conic curve. The present work aims to cope with this problem using the properties of paracatadioptric line projection .

5.2.2 Properties of a Set of Paracatadioptric Line Images

A conic curve has 5 DOF and it can be represented by a symmetric matrix $\hat{\Omega}$ (equation 5.4), or by a point $\hat{\omega} = (a, b, c, d, e, f)^t$ in P^5 [59]. Consider the line image $\hat{\Omega}$ provided in equation 5.4. Replacing $\bar{\Omega}$ and \mathbf{H}_c by the results of equations 5.3 and 5.5 yields

$$\begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix} = \begin{bmatrix} -\frac{n_z^2}{\alpha^2 f_c^2} \\ \frac{n_z^2 s_k}{\alpha f_c^3} \\ -\frac{n_z^2}{f_c^2} \left(\frac{s_k^2}{f_c^2} + \alpha^2 \right) \\ \frac{n_x n_z}{\alpha f_c} - \frac{n_z^2 (\alpha s_k c_y - f_c c_x)}{\alpha^2 f_c^3} \\ \frac{\alpha n_y n_z}{f_c} + \frac{\alpha^2 n_z^2 c_y - s_k n_x n_z}{f_c^2} + \frac{s_k n_z^2 (\alpha s_k c_y - f_c c_x)}{\alpha f_c^4} \\ \frac{n_z^2 - \alpha c_y n_y n_z}{f_c} - \frac{\alpha^3 c_y^2 n_z^2 - n_x n_z (\alpha s_k c_y - f_c c_x)}{\alpha f_c^2} - \frac{n_z^2 (\alpha s_k c_y - f_c c_x^2)}{\alpha^2 f_c^4} \end{bmatrix}$$

The paracatadioptric image of a line depends on the intrinsic parameters of the system and on the orientation of the 3D plane Π (see Fig. 5.1). After some algebraic manipulation the previous result can be rewritten in the form of equation 5.7. If the calibration is known then the conic curve $\hat{\omega}$ is only described by parameters a , d and e . These three parameters encode the scale information and the orientation of plane Π containing the imaged line. Considering that conic $\hat{\omega}$ has 5 DOF, we may conclude that 3 DOF depend on the parabolic system parameters, and the remaining 2 DOF are related with the line that is projected.

$$\hat{\omega} = \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix} = \begin{bmatrix} a \\ -\frac{\alpha s_k}{f_c} a \\ \left(\frac{\alpha^2 s_k^2}{f_c^2} + \alpha^4 \right) a \\ d \\ e \\ -\alpha^2 f_c^2 a - c_x d - c_y e \end{bmatrix} \quad (5.7)$$

Assume K lines in the scene that are projected into K conic curves in the

5. The Paracatadioptric Camera

paracatadioptric image plane ($K \geq 3$). These conic curves can be represented by points of P^5 as shown in equation 5.8.

$$\hat{\omega}_i = (a_i, b_i, c_i, d_i, e_i, f_i)^t, \quad i=1,2,3 \dots K \quad (5.8)$$

Consider the result of equation 5.7. Notice that

$$\begin{aligned} \frac{b_1}{a_1} &= \frac{b_2}{a_2} = \frac{b_3}{a_3} = \dots = \frac{b_K}{a_K} = -\frac{\alpha s_k}{f_c} \\ \frac{c_1}{a_1} &= \frac{c_2}{a_2} = \frac{c_3}{a_3} = \dots = \frac{c_K}{a_K} = \frac{\alpha^2 s_k^2}{f_c^2} + \alpha^4 \end{aligned}$$

From the first expression it results that $\eta_i = 0$ for $i = 2, 3 \dots K$ with η_i provided by equation 5.9. Moreover, using the second expression in a similar manner, comes that $\chi_i = 0$ for $i = 2, 3 \dots K$ where χ_i is given by equation 5.10.

$$\eta_i = a_1 b_i - a_i b_1, \quad i=2,3 \dots K \quad (5.9)$$

$$\chi_i = a_1 c_i - a_i c_1, \quad i=2,3 \dots K \quad (5.10)$$

From equation 5.7 it results that each line image $\hat{\omega}_i$ must verify $\alpha^2 f_c^2 a_i + c_x d_i + c_y e_i + f_i = 0$. Consider the conic curves $\hat{\omega}_1, \hat{\omega}_2$ and $\hat{\omega}_3$, which are the first three elements of the set of line images. $\alpha^2 f_c^2, c_x$ and c_y can be determined as follows

$$\begin{bmatrix} \alpha^2 f_c^2 \\ c_x \\ c_y \end{bmatrix} = - \underbrace{\begin{bmatrix} a_1 & d_1 & e_1 \\ a_2 & d_2 & e_2 \\ a_3 & d_3 & e_3 \end{bmatrix}}_{\Phi}^{-1} \underbrace{\begin{bmatrix} f_1 \\ f_2 \\ f_3 \end{bmatrix}}_{\Gamma}$$

If $K > 3$ then each conic curve $\hat{\omega}_i$ with $i = 4 \dots K$ must verify the constraint $\nu_i = 0$ (equation 5.12).

$$\nu_i = \begin{bmatrix} a_i & d_i & e_i & f_i \end{bmatrix} \cdot \begin{bmatrix} -\Phi^{-1}\Gamma \\ 1 \end{bmatrix}, \quad i=4 \dots K \quad (5.12)$$

It is clear that if a set of K conic curves corresponds to the paracatadioptric projection of K lines, then η_i, χ_i and ν_i , provided in equations 5.9, 5.10 and 5.12, must be equal to zero. We have derived $3K - 5$ independent conditions which

5.2. Paracatadioptric Camera Calibration Using Lines

are necessary for the conic curves to be paracatadioptric line images. However it has not been proved that these conditions are also sufficient. By sufficient we mean that, if a certain set of conic curves verifies these conditions then it can be the paracatadioptric projection of a set of lines.

Consider the uncalibrated image of K lines that are mapped in the same number of conics. Since each conic has 5 DOF then a set of K conics has a total of $5K$ DOF. Each line introduces 2 unknowns (DOF), which correspond to the orientation of the associated plane Π (see Fig. 5.1). Moreover the 5 parameters of matrix \mathbf{H}_c are also unknown (equation 5.5). Thus there are a total of $2K + 5$ unknowns (DOF). Since $5K > 2K + 5$ then it is obvious that there are sets of conic curves that can never be the paracatadioptric projection of lines. The conics that can correspond to the image of the lines lie in a subspace of dimension $2K + 5$. This means that there are $3K - 5$ independent constraints, which proves the sufficiency of the conditions derived above.

5.2.3 Estimation of a Set of K Paracatadioptric Line Images

Assume that we have a paracatadioptric image of K lines in the scene. Each line is projected in a conic curve $\hat{\omega}_i$ (equation 5.8). The goal is to correctly estimate the set of conic curves knowing neither the system calibration nor the position of the lines in the scene.

Conic Fitting Based on Algebraic Distances

Consider the image points $\hat{\mathbf{x}}_j^i = (\hat{x}_j, \hat{y}_j)^t$ with $j = 1, 2 \dots M_i$ and $M_i \geq 5$, lying on conic $\hat{\omega}_i$. The LMS method, discussed on chapter 4, fits the data points by the conic which minimizes the sum of the square of the algebraic distances. The function to be minimized is $\phi(\hat{\omega}_i) = \hat{\omega}_i^t \mathbf{A}_i^t \mathbf{A}_i \hat{\omega}_i$ where \mathbf{A}_i is the design matrix

$$\mathbf{A}_i = \begin{bmatrix} \hat{x}_1^2 & 2\hat{x}_1\hat{y}_1 & \hat{y}_1^2 & 2\hat{x}_1 & 2\hat{y}_1 & 1 \\ \hat{x}_2^2 & 2\hat{x}_2\hat{y}_2 & \hat{y}_2^2 & 2\hat{x}_2 & 2\hat{y}_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \hat{x}_{M_i}^2 & 2\hat{x}_{M_i}\hat{y}_{M_i} & \hat{y}_{M_i}^2 & 2\hat{x}_{M_i} & 2\hat{y}_{M_i} & 1 \end{bmatrix} \quad (5.13)$$

Considering the entire set of conic curves $\mathbf{p} = (\hat{\omega}_1^t, \hat{\omega}_2^t \dots \hat{\omega}_K^t)^t$ the design

5. The Paracatadioptric Camera

matrix becomes

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_3 & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{A}_K \end{bmatrix} \quad (5.14)$$

And the sum of the square of the algebraic distances between the data points and the set of line images is provided by function ϵ_ϕ

$$\epsilon_\phi(\mathbf{p}) = \mathbf{p}^t \mathbf{A}^t \mathbf{A} \mathbf{p} \quad (5.15)$$

One way to estimate the set of conic curves is to find the solution \mathbf{p} that minimizes ϵ_ϕ under the constraint $\mathbf{p}^t \mathbf{p} = 1$. The minimizer is the normalized eigenvector of $\mathbf{A}^t \mathbf{A}$ corresponding to the smallest eigenvalue (see section 4.2.1). As discussed in chapter 4 the problem is that in general the conic curves corresponding to the paracatadioptric projection of lines are strongly occluded in the image. The standard conic fitting techniques do not work properly under these circumstances since the data points do not provide enough information to correctly estimate the conics. Section 5.2.2 shows that a set of K conic curves corresponds to the paracatadioptric projection of K lines, if and only if, it verifies the constraints provided by equations 5.9, 5.10 and 5.12. Our approach consists in using the necessary and sufficient conditions derived above to constrain as much as possible the search space in the conic fitting problem.

General Situation

Assume the image of K lines acquired by an uncalibrated paracatadioptric camera. Nothing is known about the parameters of matrix \mathbf{H}_c . The skew can be non null and the aspect ratio different from one. Function ϵ_ϕ provides the algebraic distance between the set of conic curves \mathbf{p} and the data points (equation 5.15). We aim minimize of the algebraic distance under the constraints $\eta_i = 0$, $\chi_i = 0$ and $\nu_i = 0$ (equations 5.9, 5.10 and 5.12). One way to achieve this goal is to find the solution \mathbf{p} which minimizes the function ϵ provided in equation 5.16.

5.2. Paracatadioptric Camera Calibration Using Lines

$$\epsilon_g(\mathbf{p}) = \epsilon_\phi(\mathbf{p}) + k \left(\sum_{i=2}^K \eta_i^2 + \sum_{i=2}^K \chi_i^2 + \sum_{i=4}^K \nu_i^2 \right) \quad (5.16)$$

The constraints are introduced as penalty terms weighted by a parameter k . The minimization of the function ϵ_g can be stated as a nonlinear least squares problem. The solution can be found using Gauss-Newton or Levenberg-Marquardt algorithms [58, 27]. Notice that the Jacobian matrix can be explicitly derived in a straightforward manner.

Skewless Images with Known Aspect Ratio

Assume that the orthographic camera is skewless and that the aspect ratio α^2 is known. Replacing s_k by 0 in equation 5.7 yields $b = 0$ and $c = \alpha^4 a$. The constraints $\eta_i = 0$ and $\chi_i = 0$ for $i = 2 \dots K$, become $b_i = 0$ and $c_i - \alpha^4 a_i = 0$ for $i = 1 \dots K$. Notice that there are two additional constraints because now two of the calibration parameters are known. The new function ϵ_s is given by equation 5.17.

$$\epsilon_s(\mathbf{p}) = \epsilon_\phi(\mathbf{p}) + k \left(\sum_{i=1}^K b_i^2 + \sum_{i=1}^K (c_i - \alpha^4 a_i)^2 + \sum_{i=4}^K \nu_i^2 \right) \quad (5.17)$$

The minimization of function ϵ_s has not a closed form solution. The minima must be determined by iterative nonlinear least squares which requires an initial estimate of \mathbf{p} [58, 27]. The iterative process is in general time consuming and may not converge correctly. We show now that if the set of lines has only three elements ($K = 3$) then the problem of minimizing function ϵ_s has a closed form solution.

Consider the partial design matrix \mathbf{A}_i provided in equation 5.13. If the camera is skewless and the aspect ratio is known then $b_i = 0$ and $c_i = \alpha^4 a_i$. Omitting the second column of the design matrix and adding the third column multiplied by α^4 to the first column yields

5. The Paracatadioptric Camera

$$\check{\mathbf{A}}_i = \begin{bmatrix} \hat{x}_1^2 + \alpha^4 \hat{y}_1^2 & 2\hat{x}_1 & 2\hat{y}_1 & 1 \\ \hat{x}_2^2 + \alpha^4 \hat{y}_2^2 & 2\hat{x}_2 & 2\hat{y}_2 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \hat{x}_{M_i}^2 + \alpha^4 \hat{y}_{M_i}^2 & 2\hat{x}_{M_i} & 2\hat{y}_{M_i} & 1 \end{bmatrix}$$

The sum of the square of the algebraic distances between the conic curve and the data points is $\check{\phi} = \check{\omega}_i^t \check{\mathbf{A}}_i^t \check{\mathbf{A}}_i \check{\omega}_i$ where $\check{\omega}_i = (a_i, d_i, e_i, f_i)^t$. A novel design matrix $\check{\mathbf{A}}$ can be obtained by replacing \mathbf{A}_i by $\check{\mathbf{A}}_i$ for $i = 1 \dots K$ in equation 5.14. Assuming $\check{\mathbf{p}} = (\check{\omega}_1^t, \check{\omega}_2^t \dots \check{\omega}_K^t)^t$ comes that the sum of the square of the algebraic distances between the set of line images and the data points is $\check{\phi}_s$

$$\check{\phi}_s(\check{\mathbf{p}}) = \check{\mathbf{p}}^t \check{\mathbf{A}}^t \check{\mathbf{A}} \check{\mathbf{p}} \quad (5.19)$$

Since the novel design matrix $\check{\mathbf{A}}$ implicitly encodes the constraints $b_i = 0$ and $c_i - \alpha^4 a_i = 0$, function ϵ_s can be rewritten as

$$\check{\epsilon}_s(\check{\mathbf{p}}) = \check{\phi}_s(\check{\mathbf{p}}) + k \sum_{i=4}^K \nu_i^2 \quad (5.20)$$

As discussed in section 4.2.1 the eigenvector corresponding to the smallest eigenvalue of matrix $\check{\mathbf{A}}^t \check{\mathbf{A}}$ is the solution $\check{\mathbf{p}}$ which minimizes $\check{\phi}$ under the constraint $\check{\mathbf{p}}^t \check{\mathbf{p}} = 1$. Notice that the estimated set of conic curves also verifies the constraints $b_i = 0$ and $c_i - \alpha^4 a_i = 0$. If $K = 3$ then the second term of equation 5.20 disappears and the problem becomes closed form. Whenever $K > 3$ the minimization of function $\check{\epsilon}_s$ is a nonlinear least squares problem which must be solved using iterative procedures. However, even in these circumstances, the eigenvector solution is in general quite accurate. In this case the conditions of equation 5.12 are neglected and the search space is not fully constrained. Nevertheless it is constrained enough to provide good results.

5.2.4 Performance Evaluation Using Simulated Images

In this section we use simulated images to compare and evaluate the robustness of our approach.

5.2. Paracatadioptric Camera Calibration Using Lines

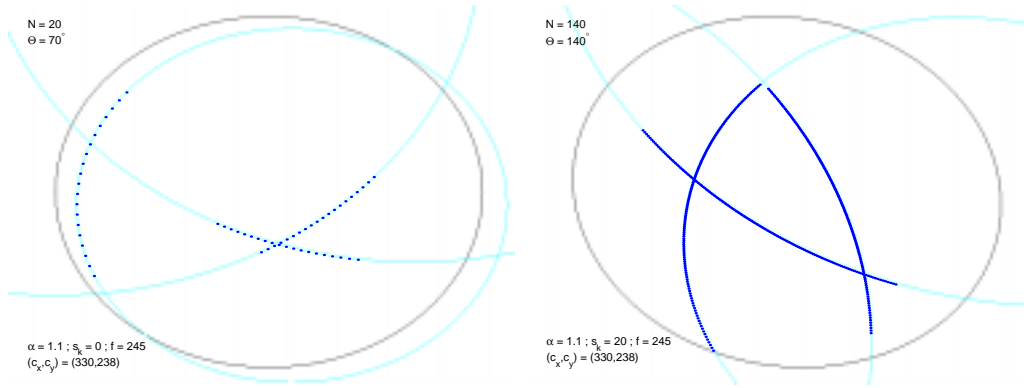


Figure 5.2: Simulated 480×640 images of three randomly generated lines

Simulation Scheme

Assume a paracatadioptric camera with a field of view (FOV) of 180° , corresponding to a full hemisphere, and predefined intrinsic parameters. The image of a set of K lines is generated as follows. As depicted in Fig. 5.1 to each line in the scene corresponds a plane Π with normal \mathbf{n} . The K normals are unitary and randomly chosen from an uniform distribution in the sphere. Each normal defines a plane that intersects the unit sphere in a great circle. Notice that half of the great circle is within the camera field of view (the FOV is 180°). An angle θ , less or equal to the FOV, is chosen to be the amplitude of the arc that is actually visible in the paracatadioptric image. The arc is randomly and uniformly positioned along the part of the great circle which is within the FOV. The visible arc is uniformly sampled by a fixed number N of sample points. The each sample point corresponds a projective ray \mathbf{x} . The sample rays are projected using formula 5.1 and transformed using 5.2 with the chosen intrinsic parameters. Two dimensional gaussian noise with zero mean and standard deviation σ is added to each image point $\hat{\mathbf{x}}$. Fig. 5.2 depicts two simulated images of 3 randomly generated lines. In the left image the visible arc has an amplitude $\theta = 70^\circ$ and is sampled by 20 points. The camera intrinsic parameters appear in the bottom left corner. In the right image the visible arc is $\theta = 140^\circ$ and the number of sample points is $N = 140$. In this case the camera is not skewless. As a final remark notice that the amplitude of the visible arc is measured in the great circle where plane Π intersects the sphere, and not in the conic curve where the line is projected. In general the visible angle of the

5. The Paracatadioptric Camera

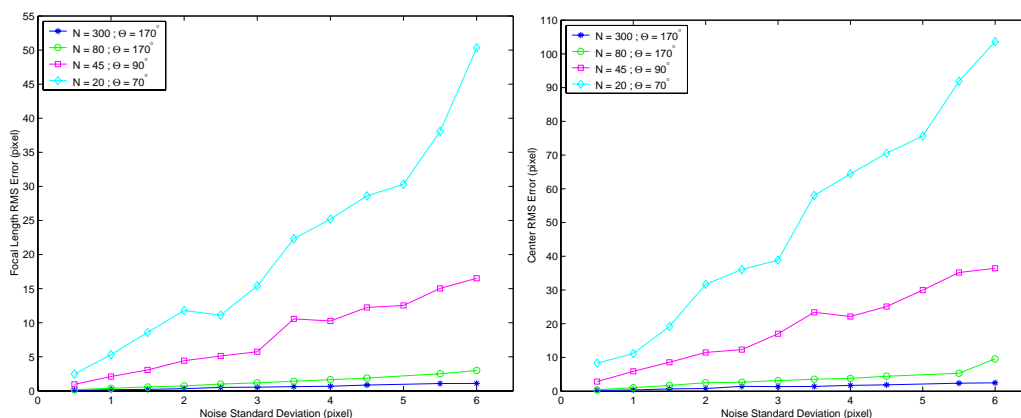


Figure 5.3: RMS error in the calibration parameters using the closed form algorithm

paracatadioptric line image is much less than θ .

Calibration of Skewless Camera with Known Aspect Ratio

Consider a parabolic camera such that the skew is 0 ($s_k = 0$), the aspect ratio is 1.21 ($\alpha = 1$), and both are assumed to be known. We wish to determine the focal length ($f_c = 245$) and the image center ($(c_x, c_y) = (330, 238)$) using the image of three lines ($K = 3$). The line images are estimated by minimizing the sum of the square of the algebraic distances $\check{\phi}_s$ (equation 5.19). As discussed in these circumstances the problem has a closed form solution. The system is calibrated using the algorithm presented in Tab. 5.1. The data points are artificially generated using the simulation scheme explained above. The left image of Fig. 5.2 is an example of a test image. The estimated calibration parameters are compared with the ground truth and the RMS error is computed over 100 runs of each experiment.

Fig. 5.3 shows the results for different choices of θ (amplitude of the visible arc) and N (number of sample points). For each choice of θ and N the standard deviation of the additive gaussian noise varies between 0.5 and 6 pixels by increments of 0.5 pixels. For $\theta = 170^\circ$ the algorithms presents an excellent performance. The decrease on the number of sample points from 300 to 80 only slightly affects the robustness to noise. Since we are only using three lines, the decrease on the amplitude of the visible arc θ and on the number of points N has

5.2. Paracatadioptric Camera Calibration Using Lines

a strong impact on the performance. Even so the calibration using arcs of 90° is still practicable. The situation of $\theta = 70^\circ$ and $N = 20$ is very extreme (left image in Fig. 5.2) leading to a bad estimation of the intrinsic parameters.

An alternative calibration approach is presented in [35]. The authors evaluate the performance of their algorithm using similar simulation conditions. A direct comparison can be made between the results presented in here and the ones presented in [35]. In general terms they estimate the conic curves by exploiting the fact that the image center must lie in the line going through the intersection points of any two line images. As discussed in section 5.2.2, this condition is necessary, but not sufficient, for a set of conic curves to be the paracatadioptric projection of lines. Since the search space is not fully constrained, they need much more than three line images to calibrate the sensor. The results presented in Fig. 5.3 are obtained using the minimum theoretical number of lines for calibration [8]. Even so, and as far as we are able to judge from the results presented in [35], the performance of our approach seems to be significantly better.

Calibration of General Paracatadioptric Systems

Artificial images are generated as explained in section 5.1. Fig. 5.2 depicts on the right one of the simulated images used during this experiment. The camera intrinsic parameters are provided in the bottom left corner. It is assumed that nothing is known about these parameters. We wish to determine the aspect ratio, skew, focal length and image center using a set of K line images.

The set of line images is estimated by minimizing the function ϵ_g provided by equation 5.16. As discussed the minimization can be stated as a nonlinear least squares problem. There is no closed form solution and the minima is found by iterative gradient descend methods such as the Gauss-Newton or the Levenberg-Marquardt methods [58, 27]. An initial estimation is needed to start the iterative minimization. The starting point is the minimizer of the sum of the square of the algebraic distances ϵ_ϕ (equation 5.15). After estimating the paracatadioptric line images the calibration parameters are computed following the steps enunciated on Tab. 5.1. The results are compared with the ground truth and the median error is computed over 100 runs. The median error is used instead of the RMS error because in some runs the iterative minimization does not converge correctly. Fig.

5. The Paracatadioptric Camera

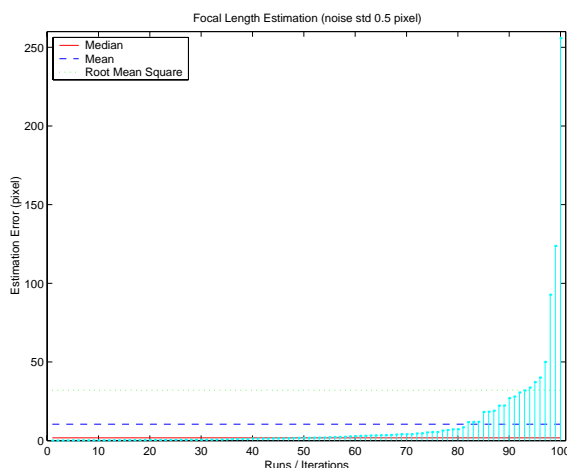


Figure 5.4: Median error, mean error and RMS error

5.4 shows the focal length estimation error, sorted in ascending order, over the 100 runs of a certain experiment. The mean, median and RMS error are depicted by the horizontal lines. As you can observe it is reasonable to argue that the median error is the one that best reflects the result of the experiment.

In this first set of experiments we aim to calibrate the system using the minimum theoretical number of lines ($K = 3$). The visible arc θ is always 170° but the number of sample points N changes ($N = 300, 170, 80$). For each choice of θ and N the standard deviation of the additive gaussian noise varies between 0.5 and 6 pixels by increments of 0.5 pixels.

The minimization of function ϵ_g is performed using iterative gradient descending methods. This can be problematic in many ways [58, 27]. The initial estimate is crucial to assure the convergence to the right solution. It is important to start the iterative process from a point as close as possible to the global minima. Moreover the objective function may have local minima and saddle points. This happens often when there is not enough information to correctly constrain the problem. By lack of information we mean small number of lines, data points strongly corrupted with noise, visible arcs with small amplitude or not sampled enough. In these circumstances the iterative minimization may not converge to the right solution. Fig. 5.5 shows the number of convergence failures over the 100 runs of each experiment. The run fails when the absolute conic $\hat{\Omega}_\infty$, determined following the

5.2. Paracatadioptric Camera Calibration Using Lines

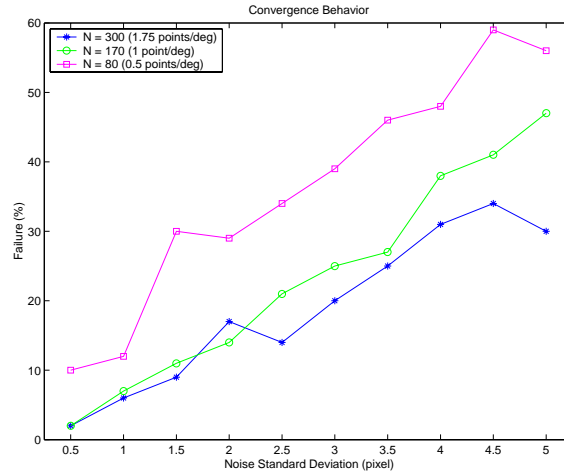


Figure 5.5: Convergence of the iterative minimization

steps in Tab. 5.1, is not positive definite and the Cholesky decomposition is not possible. This can only happen when the set of paracatadioptric line images is far from being correctly estimated. As expected the convergence is strongly affected by the noise. Moreover the decrease in the number of sample points also causes an increase in the number of failures. Notice that we are assuming three randomly generated line images. This is the minimum number of lines required to calibrate a paracatadioptric system. Thus it is natural that often the minimization process does not converge correctly.

Fig. 5.6 shows the median error in the estimation of the different calibration parameters. The performance clearly decreases when the number of sample points decreases.

The calibration results presented in Fig. 5.6 are not very impressive. However we must take into account that we are using only three line images. In this set of experiments we aim to compare the performance of the calibration algorithm when using 3, 5, 7 and 9 line images. The amplitude of the visible arc and the number of sample points are respectively $\theta = 140^\circ$ and $N = 140$ in all experiments. The median errors in estimating the calibration parameters can be observed in Fig. 5.7. The increase in the number of lines dramatically improves the robustness of the calibration.

5. The Paracatadioptric Camera

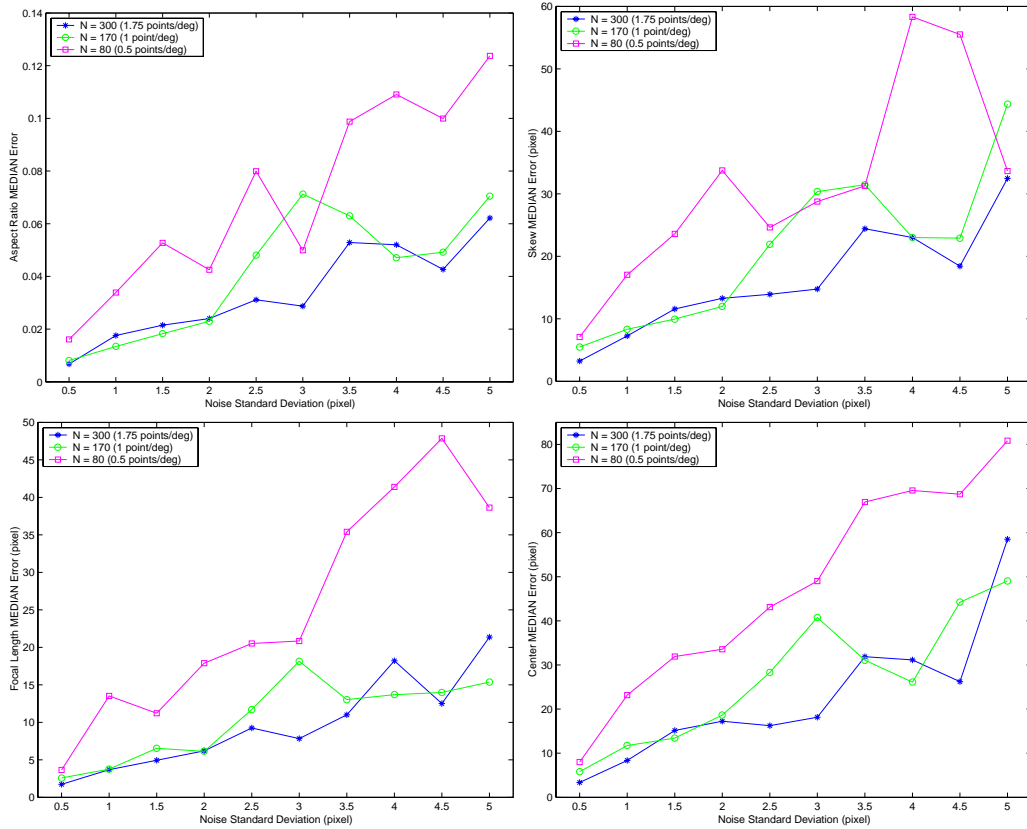


Figure 5.6: Median error in the calibration parameters using 3 lines

5.2. Paracatadioptric Camera Calibration Using Lines

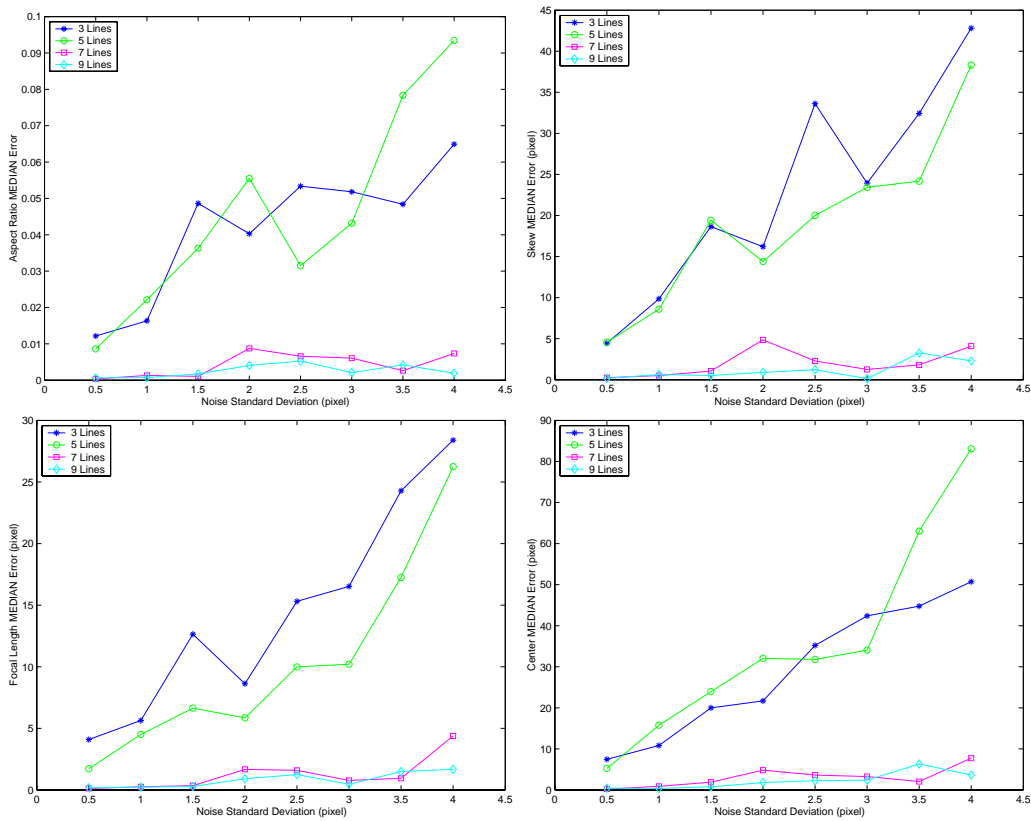


Figure 5.7: Median error in the calibration parameters using 3, 5, 7 and 9 lines

5. The Paracatadioptric Camera

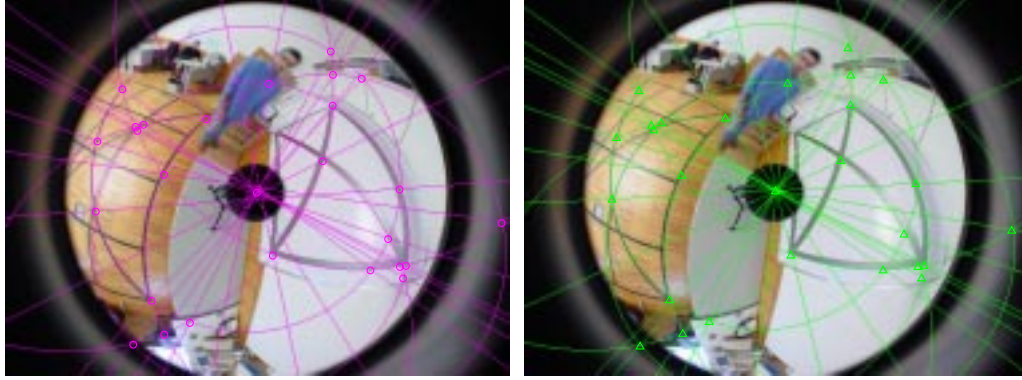


Figure 5.8: Estimating line images for calibration purposes

		3 Lines	4 Lines	5 Lines	6 Lines
α	<i>mean</i>		1.0001	0.9998	0.9996
	<i>std</i>		0.0012	0.0019	0.0015
f_c	<i>mean</i>	699.36	699.37	701.03	701.81
	<i>std</i>	17.24	16.00	13.57	10.65
s_k	<i>mean</i>		1.46	0.57	-1.95
	<i>std</i>		2.35	1.41	1.39
c_x	<i>mean</i>	1137.0	1137.6	1143.6	1147.7
	<i>std</i>	21.4	22.6	11.0	5.8
c_y	<i>mean</i>	870.90	870.66	874.36	876.64
	<i>std</i>	11.84	13.42	8.29	5.66

Table 5.2: Calibration results for different number of lines

5.2.5 Experimental Results Using Real Images

This section applies the proposed calibration approach to real paracatadioptric images. Five images were taken using a paracatadioptric camera. The image resolution is 1704×2272 and the FOV is 180° . Fig. 5.8 is one of those images where a set of lines is clearly visible. For each image we used an edge detector and selected points belonging to 6 different lines. Each one of the five images are calibrated using 3, 4, 5 and 6 lines. Tab. 5.2 presents, for each situation, the mean of the calibration results as well as the corresponding standard deviation.

In this case nothing is known about the calibration parameters. Thus the esti-

5.2. Paracatadioptric Camera Calibration Using Lines

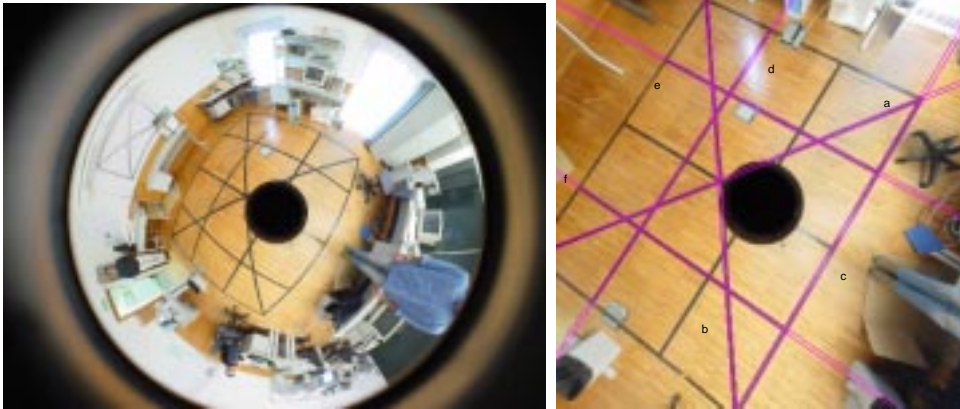


Figure 5.9: Perspective obtained by rectifying a paracatadioptric image

mation of the set of line images is performed by finding the solution which minimizes the function ϵ_g of equation 5.16. However notice that the image of the mirror boundary is close to a circle which allows us to infer that the skew is small and the aspect ratio is approximately unitary. Taking this into account the initial estimation for the iterative process is obtained using the closed form algorithm proposed at the end of section 5.2.3. The left and right images on Fig 5.8 show the initial and final estimate of the set of conic curves. The calibration results are summarized in Tab. 5.2. Notice that the estimated values for the calibration parameters are more or less the same for the different K (number of lines). The standard deviation acts as a measure of confidence. If the standard deviation takes high values then the results obtained for each image are very different and the achieved calibration is not trustable. As expected the standard deviation decreases when the number of lines increases.

To evaluate the correctness of the results we have rectified a paracatadioptric image. The obtained perspective image is exhibited on Fig. 5.9, where six pairs of parallel lines are indicated by letters. The lines were estimated using normal least squares. Consider the direction of those pairs of lines. The angle between each two directions can be determined using the corresponding vanishing points and the image of the absolute conic. The vanishing points are the intersections of the images of the parallel lines, and the absolute conic is known since the perspective image is artificially generated. The angles between each two directions were

5. The Paracatadioptric Camera

computed and the results were compared with the angles measured in the scene. The mean of the error was 0.51° and standard deviation was 0.35° .

5.3 Direct Least Square Fitting of Paracatadioptric Line Images

In general the paracatadioptric image of a line is a conic curve. The estimation of line images is an important subject for applications such as reconstruction and visual control of motion. However the estimation of the conic curves where lines are mapped is hard to accomplish. In general only a small arc of the conic is visible in the image and conventional conic fitting techniques are unable to correctly estimate the curve (section 4.2). This section shows that for a calibrated paracatadioptric system the line images can be accurately estimated by constraining the search space. A conic curve is the paracatadioptric image of a line if, and only if, the image of the circular points lie on the curve and two certain points are conjugate with respect to the conic. Considering the space of all conic curves, the line images lie in a linear subspace which depends on the system calibration. The paracatadioptric projection of a line can be estimated by fitting a conic in the subspace to the data points. The proposed approach is computationally efficient since the fitting problem can be solved by an eigensystem

5.3.1 The Necessary and Sufficient Condition

Consider the scheme of Fig. 5.1 for the mapping of a line by a paracatadioptric sensor. Plane Π , containing both the line and the effective viewpoint O , intersects the sphere in a great circle. The mapping from the sphere to plane Π_∞ is a stereographic projection. The great circle is projected into $\bar{\Omega}$ which is circle as can be verified by inspecting equation 5.3. Points in plane Π_∞ are mapped into points in the image by a collineation H_c . Notice that for the parabolic situation H_c is always an affine transformation (equation 5.2). Since an affine transformation does not change the type of conic, then the paracatadioptric image of a line $\hat{\Omega}$ is always a circle/ellipse (equation 5.4).

Consider the following points lying on plane Π_∞ :

5.3. Direct Least Square Fitting of Paracatadioptric Line Images

$$\begin{aligned}\bar{\mathbf{I}} &= (1, i, 0)^t \\ \bar{\mathbf{J}} &= (1, -i, 0)^t \\ \bar{\mathbf{G}} &= (1, 0, -i)^t \\ \bar{\mathbf{H}} &= (1, 0, i)^t\end{aligned}$$

Assume that the paracatadioptric system is calibrated and the affine transformation \mathbf{H}_c is known. The points are mapped in the paracatadioptric image plane on points:

$$\begin{aligned}\hat{\mathbf{I}} &= \mathbf{H}_c \bar{\mathbf{I}} = (i_x, i_y, i_z)^t \\ \hat{\mathbf{J}} &= \mathbf{H}_c \bar{\mathbf{J}} = (j_x, j_y, j_z)^t \\ \hat{\mathbf{G}} &= \mathbf{H}_c \bar{\mathbf{G}} = (g_x, g_y, g_z)^t \\ \hat{\mathbf{H}} &= \mathbf{H}_c \bar{\mathbf{H}} = (h_x, h_y, h_z)^t\end{aligned}$$

According to the result of proposition 3.6 (section 3.3.2) the conic curve $\hat{\Omega}$ is the paracatadioptric projection of a line if, and only if, it verifies the conditions $\hat{\mathbf{I}}^t \hat{\Omega} \hat{\mathbf{I}} = 0$, $\hat{\mathbf{J}}^t \hat{\Omega} \hat{\mathbf{J}} = 0$ and $\hat{\mathbf{G}}^t \hat{\Omega} \hat{\mathbf{H}} = 0$. The conic curve $\hat{\Omega}$ can be parameterized by a point $\hat{\omega} = (a, b, c, d, e, f)^t$ in \mathbf{P}^5 . Thus $\Upsilon \hat{\omega} = 0$, with Υ the 3×6 matrix provided in equation 5.23, is a necessary and sufficient condition for a conic $\hat{\omega}$ being a paracatadioptric line image.

$$\Upsilon = \begin{bmatrix} i_x^2 & 2i_x i_y & i_y^2 & 2i_x i_z & 2i_y i_z & i_z^2 \\ j_x^2 & 2j_x j_y & j_y^2 & 2j_x j_z & 2j_y j_z & j_z^2 \\ g_x h_x & g_x g_y + h_x h_y & g_y h_y & g_x g_z + h_x h_z & g_y g_z + h_y h_z & g_z h_z \end{bmatrix} \quad (5.23)$$

5.3.2 The Algorithm

Consider the set of image points $\hat{\mathbf{x}}_i = (\hat{x}_i, \hat{y}_i)^t$ with $i = 1, 2, \dots, M$. The goal is to fit a conic curve $\hat{\omega}$, corresponding to the paracatadioptric projection of a line, to the set of data points. The sum of the square of the algebraic distances between the curve and the image points is $\alpha = \hat{\omega}^t \mathbf{A}^t \mathbf{A} \hat{\omega}$ with \mathbf{A} the design matrix

5. The Paracatadioptric Camera

$$\mathbf{A} = \begin{bmatrix} \hat{x}_1^2 & 2\hat{x}_1\hat{y}_1 & \hat{y}_1^2 & 2\hat{x}_1 & 2\hat{y}_1 & 1 \\ \hat{x}_2^2 & 2\hat{x}_2\hat{y}_2 & \hat{y}_2^2 & 2\hat{x}_2 & 2\hat{y}_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \hat{x}_M^2 & 2\hat{x}_M\hat{y}_M & \hat{y}_M^2 & 2\hat{x}_M & 2\hat{y}_M & 1 \end{bmatrix}$$

The conic curve $\hat{\omega}$, represented by a point in \mathbf{P}^5 , has 5 degrees of freedom (DOF). Neglecting the scale factor the space of all conics has five dimensions. The conic fitting algorithms discussed in section 4.2 search the entire space for the conic that best fits the data accordingly to a certain criteria. However not all conics can be the paracatadioptric image of a line. The paracatadioptric line image $\hat{\omega}$ must lie in the null space of matrix Υ (equation 5.23). The null space of Υ is a linear subspace in the space of all conic curves. Our approach fits the data by the conic curve, lying in this subspace, that minimizes the algebraic distance to the image points. Consider the singular value decomposition of matrix Υ .

$$\Upsilon = \mathbf{U}\mathbf{S}\mathbf{V}^t$$

Matrices \mathbf{U} , \mathbf{S} and \mathbf{V} have respectively dimension 3×3 , 3×6 and 6×6 . \mathbf{V} is full rank and orthonormal ($\mathbf{V}^{-1} = \mathbf{V}^t$). The three last columns of \mathbf{V} are an orthonormal basis of the null space of Υ [58, 38]. Consider the change on the base of representation $\hat{\omega}_{\mathbf{v}} = \mathbf{V}\hat{\omega}$. If $\hat{\omega}$ belongs to the null space of matrix Υ , then the corresponding $\hat{\omega}_{\mathbf{v}}$ has the following structure

$$\hat{\omega}_{\mathbf{v}} = (0, 0, 0, \underbrace{d_v, e_v, f_v}_{\rho})^t \quad (5.26)$$

The algebraic distance between conic $\hat{\omega}$ and the data points is $\alpha = \hat{\omega}^t \mathbf{A}^t \mathbf{A} \hat{\omega}$. Rewriting the algebraic distance in terms of the new coordinates arises $\alpha = \hat{\omega}_{\mathbf{v}}^t \mathbf{V} \mathbf{A}^t \mathbf{A} \mathbf{V}^t \hat{\omega}_{\mathbf{v}}$. Taking into account the structure of $\hat{\omega}_{\mathbf{v}}$ (equation 5.26) comes that $\alpha = \rho^t \mathbf{\Lambda}^t \mathbf{\Lambda} \rho$ with $\mathbf{\Lambda}$ the bottom right 3×3 sub matrix of $\mathbf{V} \mathbf{A}^t \mathbf{A} \mathbf{V}^t$. We aim to determine the solution ρ which minimizes the algebraic distance d under the constraint $\rho^t \rho = 1$. The objective function is

$$\phi_{catparb}(\rho, \lambda) = \rho^t \mathbf{\Lambda}^t \mathbf{\Lambda} \rho + \lambda(\rho^t \rho - 1) \quad (5.27)$$

5.3. Direct Least Square Fitting of Paracatadioptric Line Images

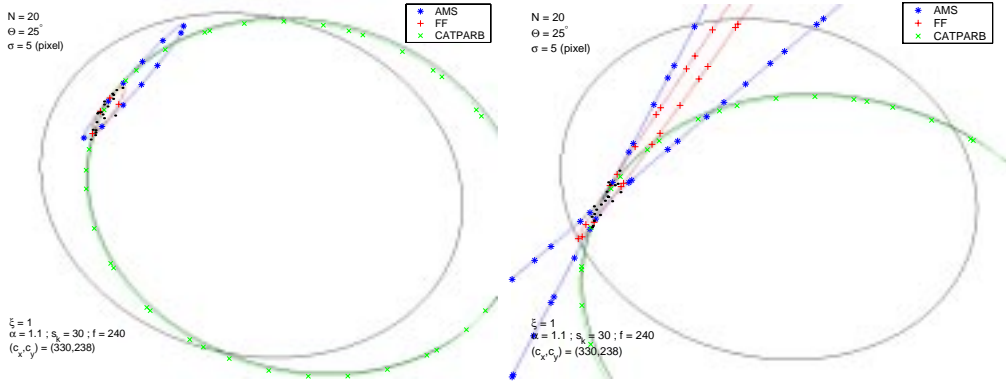


Figure 5.10: Estimating paracatadioptric line image using different methods

The minima of the objective function $\phi_{catparb}$ is the eigenvector of matrix $\Lambda^t \Lambda$ corresponding to the smallest eigenvalue. The final conic $\hat{\omega}$ is computed by replacing ρ in equation 5.27 and making $\hat{\omega} = \mathbf{V}^t \hat{\omega}_v$.

5.3.3 Experiments

The present section evaluates the performance of the proposed CATPARB algorithm. Simulated images, as the ones exhibited in Fig. 5.10, are used to characterize the method and compare its performance with other approaches. The artificial data is generated using the simulation scheme explained in section 5.2.4. Experiments using real images are also presented.

CATPARB Versus Standard Conic Fitting Methods

Several standard conic fitting techniques have been introduced in section 4.2. The present section compares the AMS and FF methods with the proposed CATPARB algorithm.

Consider the graphic on Fig. 5.11 which compares the performance of the three methods. The data points are artificially generated using the simulation scheme described above. An arc with an amplitude 80° is uniformly sampled by 40 points. Each method fits a conic curve to the data points. The estimated conic is compared with the ground truth and the RMS error in the principal points is computed over 100 runs of each experiment. Both Fig. 5.10 and 5.11 show that

5. The Paracatadioptric Camera

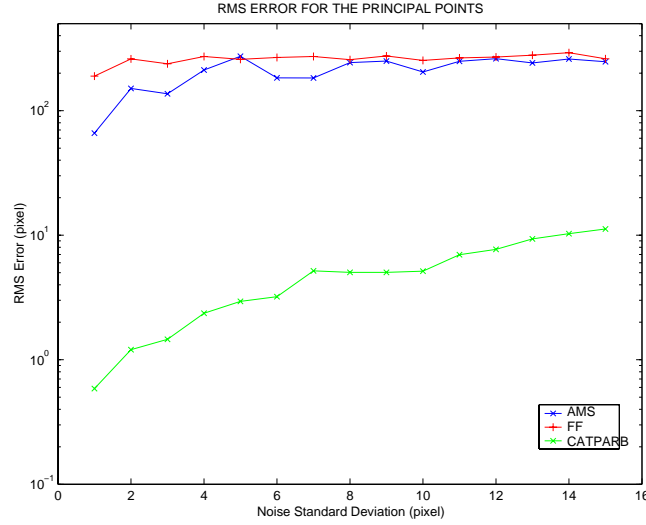


Figure 5.11: Comparing the performance of AMS, FF and CATPARB methods ($N = 40, \theta = 80^\circ$)

the estimation results dramatically improve when using the CATPARB algorithm.

CATPARB Versus Perspective Rectification

Fig. 5.11 shows that CATPARB performs much better than AMS and FF methods in estimating the conic locus where a line is mapped by a paracatadioptric sensor. Notice however that this comparison is not entirely fair. While the FF and AMS are generic methods to fit a conic curve to image points, the CATPARB algorithm uses information about the sensor geometry and calibration to perform the estimation. CATPARB is a specific method which requires both the data points $\hat{\mathbf{x}}_i$ and the calibration matrix \mathbf{H}_c to estimate the paracatadioptric line image.

Since the system calibration is known then the line can be determined in a straightforward manner by performing a perspective rectification of the data points. Consider the calibration matrix \mathbf{H}_c (equation 5.2), the inverse function \tilde{h}^{-1} (equation 2.9) where ξ is made unitary ($\xi = 1$), and the data points $\hat{\mathbf{x}}_i$ lying on the paracatadioptric line image $\hat{\Omega}$. The equation below computes the rectified data points \mathbf{x}_i

$$\mathbf{x}_i = \tilde{h}^{-1}(\mathbf{H}_c^{-1}\hat{\mathbf{x}}_i) \quad i=1 \dots N \quad (5.28)$$

5.3. Direct Least Square Fitting of Paracatadioptric Line Images

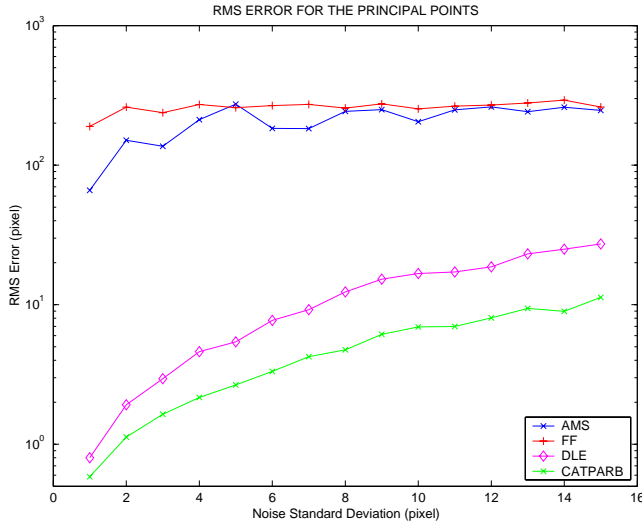


Figure 5.12: Comparing the performance of AMS, FF and CATPARB methods with the direct line estimation (DLE) after perspective rectification ($N = 40$, $\theta = 80^\circ$)

Assume that $\mathbf{n} = (n_x, n_y, n_z)^t$ is the normal to the plane Π containing the line in the scene which is imaged in $\hat{\Omega}$ (see Fig. 5.1). If point $\hat{\mathbf{x}}_i$ lies in the conic locus $\hat{\Omega}$, then the rectified point \mathbf{x}_i lies on line \mathbf{n} in the conventional perspective plane. Thus, given the set of rectified data points \mathbf{x}_i , $i = 1 \dots N$, we can estimate line \mathbf{n} using normal least squares (section 4.2.1). The solution is obtained by computing the eigenvector of the design matrix \mathbf{A} (equation 5.29) corresponding to the smallest eigenvalue. The conic locus $\hat{\Omega}$ in the image plane is computed from $\mathbf{n} = (n_x, n_y, n_z)^t$ and \mathbf{H}_c using the relations established in equations 5.3 and 5.4.

$$\mathbf{A} = \begin{bmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ \vdots & \vdots & \vdots \\ x_N & y_N & 1 \end{bmatrix} \quad (5.29)$$

Fig. 5.12 compares the performance of AMS, FF and CATPARB algorithms with the direct line estimation (DLE) after perspective rectification. The simulation scheme is similar to the one used on previous section and the metric is the

5. The Paracatadioptric Camera

RMS error on the principal points of the estimated paracatadioptric line image. It can be observed that the DLE method performs much better than the standard conic fitting techniques (AMS and FF). This is explained by the fact that DLE uses, not only the data points, but also information about the image formation process and sensor calibration. However its performance is clearly worse than the one of CATPARB method.

As explained in section 4.2.2 it is reasonable to assume that the noise in the image points $\hat{\mathbf{x}}_i = (\hat{x}_i, \hat{y}_i)^t$ is gaussian, two dimensional and with zero mean. It is also reasonable to assume that the error is equal in both directions and uncorrelated. Thus the noise covariance matrix has the form $\sigma^2 \mathbf{I}$ with \mathbf{I} the 2×2 identity matrix. Consider the rectified point \mathbf{x}_i and the line \mathbf{n} , both lying in the conventional perspective plane. The algebraic distance between the point and the line is $\alpha_i = \mathbf{n}^t \mathbf{x}_i$. Replacing \mathbf{x}_i by the result of equation 5.28 and assuming $\mathbf{H}_c = \mathbf{I}$ yields $\alpha_i = \mathbf{n}^t \cdot \tilde{h}^{-1}(\hat{\mathbf{x}}_i)$. Propagating the variance of the image point $\hat{\mathbf{x}}_i$ comes that the noise variance in the algebraic distance is

$$\sigma_i^2 = 4 \frac{(n_x^2 + n_y^2)(1 + (\hat{x}_i^2 + \hat{y}_i^2)) + 2(n_x^2 + n_y^2)(\hat{x}_i^2 + \hat{y}_i^2) + 8n_x n_y \hat{x}_i \hat{y}_i}{(1 - \hat{x}_i^2 - \hat{y}_i^2)^4} \sigma^2 \quad (5.30)$$

The least square estimator computes the line \mathbf{n} which minimizes the sum of the squares of the algebraic distances α_i ($i = 1 \dots N$). The estimation is optimal in terms of minimum covariance if the noise in the algebraic distances α_i has always the same variance and is statistically independent [55, 58]. From equation 5.30 comes that the variance σ_i is a function of the coordinates of the original image point $\hat{\mathbf{x}}_i$. Thus the variance of the algebraic distances α_i is not constant and the line estimation using least squares is statistically biased [44]. The effects of the statistical bias are much stronger in the DLE method than in the CATPARB algorithm, which explains the poorer performance of the former (Fig. 5.12).

Pushing CATPARB to the Limit

The graphic of Fig. 5.13 shows the behavior of the proposed approach in the presence of increasing noise for different values of N and θ . As expected the performance is worse when the number of samples and/or the amplitude of the

5.3. Direct Least Square Fitting of Paracatadioptric Line Images

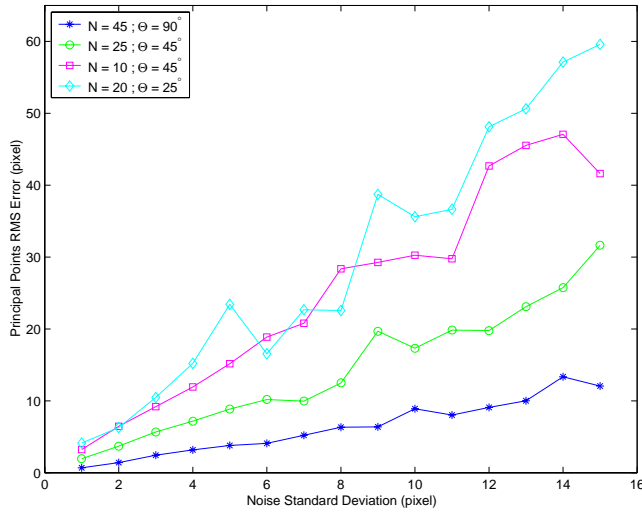


Figure 5.13: Characterization of the performance of the CATPARB algorithm.

visible arc decrease. The results depicted on Fig. 5.13 provide a general idea of the robustness of the CATPARB algorithm.

Experiments with Real Images

This section presents some results in estimating lines in real paracatadioptric images. The image resolution is 1704×2272 and the sensor is calibrated using the method proposed in section 5.2.

Fig. 5.14 depicts one of our test images. The conic curve where a line is projected has only 2 independent degrees of freedom (section 3.3.2). Thus two image points are enough to correctly determine a paracatadioptric line image. We have selected by hand two points lying on the conic locus where a certain line in the scene is projected. The estimation results using the proposed CATPARB algorithm can be observed in Fig. 5.14. The selected points are marked using the blue crosses. Notice that in general a conic curve can only be estimated using 5 or more data points.

Consider the paracatadioptric image exhibited in Fig. 5.15 with four pairs of parallel lines denoted by a, b, c, d . According to proposition 3.1, the polar of the image center with respect to the conic locus where the line is mapped, is the horizon of the plane Π containing the original 3D line and the effective viewpoint

5. The Paracatadioptric Camera



Figure 5.14: Estimating lines in a real paracatadioptric image using only two points



Figure 5.15: Estimating the angle between pairs of parallel lines

5.3. Direct Least Square Fitting of Paracatadioptric Line Images

	G. Truth	Error CATPARB	Error DLE
a-b	90°	0.8538°	0.3442°
a-c	30°	2.0504°	6.6411°
a-d	30°	1.8212°	0.4977°
b-c	60°	2.9888°	2.4049°
b-d	60°	1.0459°	13.7148°
c-d	60°	3.8717°	16.0897°
<i>median</i>		1.9358°	4.5230°
<i>mean</i>		2.1053°	6.6154°
<i>std</i>		1.1553°	6.8505°

Table 5.3: Recovering angles between pairs of parallel lines

(see Fig. 5.1). Moreover if two imaged lines are parallel then the intersection of the corresponding horizons is the vanishing point of their common direction (corollary 3.5).

The eight line images, corresponding to the four pairs of parallel lines, are estimated using the CATPARB algorithm (see Fig. 5.15). The vanishing point of each pair is determined in a straightforward manner using the results of corollary 3.3. Since the calibration matrix \mathbf{H}_c is known then the image of the absolute conic can be computed making $\hat{\Omega}_\infty = \mathbf{H}_c^{-t} \mathbf{H}_c^{-1}$. The estimation of the angles between the pairs of parallel lines from the vanishing points and the absolute conic is trivial [39, 59]. Tab. 5.3 shows the errors in estimating these angles.

There is an alternative approach to estimate the angles between the pairs of parallel lines. Instead of applying the CATPARB method and using the geometric relations derived in chapter 3, we can perform the perspective rectification of the image points, estimate the lines using normal linear least squares and compute the angles using standard projective relations. The estimation errors are shown in the last column of Tab. 5.3. As expected estimating the lines directly in the paracatadioptric plane presents better results. We may conclude that the bias introduced by the perspective rectification has a strong impact on the performance of the DLE method.

5.4 Closure

This chapter presents an effective way to calibrate paracatadioptric cameras using the image of lines in general position. In chapter 3 we proved that a line is mapped into a conic and that any central catadioptric system can be fully calibrated from a minimum of three line images in general position. However in chapter 4 we realized that the accurate estimation of lines using image points is hard to accomplish. This chapter proposes a method to overcome this problem when using parabolic systems. If it is true that any line is imaged in a conic locus, it is not true that any conic curve is the paracatadioptric image of a line. The necessary and sufficient conditions that must be verified by a set of conic curves to be the image of a set of lines are derived. These conditions are used to constrain the search space and accurately estimate the set of conic curves required to calibrate the paracatadioptric sensor. If the camera is skewless and the aspect ratio is known then the conic fitting problem is solved naturally by an eigensystem. Otherwise the estimation is performed using non-linear optimization techniques. Experimental results show that the proposed calibration method performs much better than the ones appearing in the literature [35, 45, 74].

The second contribution in this chapter is the CATPARB algorithm to estimate lines images from data points measured in the calibrated paracatadioptric plane. It has been proved in chapter 3 that a conic curve is the paracatadioptric image of a line if, and only if, it goes through points $\hat{\mathbf{I}}_\infty$, $\hat{\mathbf{J}}_\infty$ and points $\hat{\mathbf{G}}$, $\hat{\mathbf{H}}$ are harmonic conjugate with relation to it (proposition 3.6). This provides three necessary and sufficient conditions which define a linear subspace in the space of all conic curves. The line image is estimated within this subspace by solving an eigensystem. The method is accurate, robust and computationally efficient. Experimental results show that this approach performs much better than estimating the lines using perspective rectification as is often done in robotic applications [64]. The estimation after perspective rectification is statistically biased [44] which strongly affects the results.

Chapter 6

A General Framework for Selecting the World Coordinate System

The imaging process can be interpreted as a mapping from points in 3D space into points in the 2D image plane. The image formation is a transformation from \mathbf{R}^3 to \mathbf{R}^2 . Cartesian coordinate systems are typically used to reference points both in space and in the image plane. The transformation is non-injective and implies loss of information. The relationship between position and velocity in the 3D space and position and velocity in the image are in general complex, difficult and non-linear. This chapter shows that the choice of the coordinate system to reference points in the 3D space is important. The intrinsic nature of image formation process is kept unchanged but the mathematical relationship between the world and the image becomes simpler and more intuitive. This can help not only the understanding of the imaging process but also the development of new algorithms and applications.

The first part of the chapter focuses on static imaging systems that include both perspective cameras and central catadioptric systems. A general framework to describe the mapping from 3D points to 2D points in the image plane is presented. The mathematical expression of this global mapping depends on the coordinate system used to reference points in the scene. A systematic approach to select the most suitable world coordinate system is presented and discussed. Differential constraints are defined to enable the choice of a 3D reference frame. Coordinate transformations satisfying these differential constraints bring advantageous

6. A General Framework for Selecting the World Coordinate System

properties when mapping 3D space velocities into 2D image velocities. One such coordinate transformation is described for the case of the perspective camera and then generalized for central catadioptric image formation. The coordinate transformation does not imply that new information is available in the images. Instead the geometric transformations are represented in a common and more compact framework, enabling newer insights into the image formation process. Examples and applications that benefit from an adequate choice of the world coordinate system are presented and discussed.

The second part of the chapter applies the derived mathematical framework to active tracking of moving targets [2]. For this purpose it is assumed that the imaging sensor is mounted on a moving platform. The goal of the tracking application is to control the motion of the platform in such a way that the position of the target in the image plane is kept constant. Three different cases are considered: a perspective camera with translational motion in the XY plane, a perspective camera with rotational pan and tilt motion and a parabolic omnidirectional camera with a rotational degree of freedom around the Z axis. The platforms considered in this work have less than 3 degrees of freedom (DOF). For the purpose of controlling the constrained 3D motion of these robots it is not necessary to determine the full pose of the target. It is assumed that target motion is characterized by the 3D position and velocity of the corresponding mass center in an inertial reference frame. It is also assumed that the position of each degree of freedom is known (possibly via an encoder).

For the visual control of motion the relationship between motion in the scene and motion in the image must be established. The image motion depends both on target and camera 3D motion. The mathematical expression of the global mapping depends on the world coordinates used to reference points in the scene. General criteria to select suitable coordinate systems are discussed. Adequate choices are presented for each type of platform. The derived mathematical framework is used to establish the position and velocity relationships between target 3D motion, camera motion and image motion. The expressions obtained are used to implement image based active visual tracking. Simplifications of the equations obtained (to decouple the degrees of freedom of the pan and tilt vision system) are discussed.

6.1. World Coordinates for Static Imaging Systems

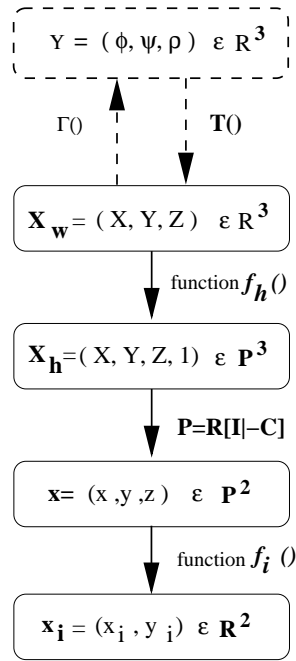


Figure 6.1: Mapping in central projection systems

6.1 World Coordinates for Static Imaging Systems

The two first sections refer to static central projection vision systems. Examples of such systems are the perspective camera and catadioptric systems that verify the fixed viewpoint constraint [5]. The image acquisition process maps points from the 3D space into the 2D image plane. Image formation performs a transformation from \mathbb{R}^3 to \mathbb{R}^2 that can be denoted by F . A generic framework to illustrate the transformation F is proposed. This framework is general to both conventional perspective cameras and central projection catadioptric systems. It is desirable that F be as simple as possible and as compact as possible. This can be achieved by selecting a specific coordinate systems to reference the world points. General criteria to select the world coordinate system are presented and discussed. Advantages of using different world coordinate systems to change the format of the F mapping are presented.

6. A General Framework for Selecting the World Coordinate System

6.1.1 Mapping Points from the 3D Space in the 2D Image Plane

Fig. 6.1 depicts a generic framework to illustrate the transformation \mathbf{F} from \mathbf{R}^3 in \mathbf{R}^2 performed by a central projection vision system. The schematic of Fig. 6.1 is an adaptation of the one of Fig. 2.3.

$\mathbf{X}_w = (X, Y, Z)^t$ is a vector with the Cartesian 3D coordinates of a point in space. The domain of transformation is the set \aleph of visible points in the world with $\aleph \subset \mathbf{R}^3$. Function \mathbf{f}_h maps \mathbf{R}^3 into the projective space \mathbf{P}^3 . It is a non-injective and surjective function transforming $\mathbf{X}_w = (X, Y, Z)^t$ in $\mathbf{X}_h = (X, Y, Z, 1)^t$ that are the homogeneous world point coordinates. \mathbf{P} is an arbitrary 3×4 homogeneous matrix with rank 3. It represents a general projective transformation performing a linear mapping of \mathbf{P}^3 into the projective plane \mathbf{P}^2 ($\mathbf{x} = \mathbf{P}\mathbf{X}_h$). The rank 3 requirement is due to the fact that if the rank is less than 3 then the range of the matrix will be a line or a point and not the whole plane. The rank 3 requirement guarantees that the transformation is surjective. In the case of \mathbf{P} being a camera model it can be written as $\mathbf{P} = \mathbf{R}[\mathbf{I} \mid -\mathbf{C}]$ where \mathbf{I} is a 3×3 identity matrix, and \mathbf{R} the rotation matrix between camera and world coordinate systems and \mathbf{C} the projection center in world coordinates [39]. If nothing is stated we will assume $\mathbf{P} = [\mathbf{I} \mid \mathbf{0}]$. Function \mathbf{f}_i transforms coordinates in the projective plane $\mathbf{x} = (x, y, z)^t$ into Cartesian coordinates in the image plane $\mathbf{x}_i = (x_i, y_i)^t$. It is a non-injective, surjective function of \mathbf{P}^2 in \mathbf{R}^2 that maps projective rays in the world into points in the image. For a conventional perspective camera with matrix of intrinsic parameters $\mathbf{K}_c = \mathbf{I}$ comes that $\mathbf{x}_i = \mathbf{f}_i(\mathbf{x}) \Leftrightarrow (x_i, y_i) = (\frac{x}{z}, \frac{y}{z})$. As it will be shown latter, for the case of general central catadioptric systems function \mathbf{f}_i encodes the non-linear transformation \tilde{h} introduced in chapter 2.

The transformation \mathbf{F} maps 3D world points into 2D points in the image. Points in the scene are represented using standard cartesian coordinates. However a different coordinate system can be used to reference points in the 3D world space. Assume that $\Upsilon = (\phi, \psi, \rho)^t$ are point coordinates in the new reference frame and that $\mathbf{X}_w = \mathbf{T}(\Upsilon)$ where \mathbf{T} is a bijective function from \mathbf{R}^3 in \mathbf{R}^3 . The transformation \mathbf{F} , mapping 3D world points Υ in image points \mathbf{x}_i (see equation 6.1), can be written as the composition of equation 6.2.

$$\mathbf{x}_i = \mathbf{F}(\Upsilon) \tag{6.1}$$

6.1. World Coordinates for Static Imaging Systems

$$\mathbf{F}(\Upsilon) = \mathbf{f}_i(\mathbf{P}\mathbf{f}_h(\mathbf{T}(\Upsilon))) \quad (6.2)$$

Equation 6.3, obtained by differentiating equation 6.1 with respect to time, establishes the relationship between velocity in 3D space $\dot{\Upsilon} = (\dot{\phi}, \dot{\psi}, \dot{\rho})^t$ and velocity in image $\dot{\mathbf{x}}_i = (\dot{x}_i, \dot{y}_i)^t$. $\dot{\mathbf{x}}_i$ and $\dot{\Upsilon}$ are related by the jacobian matrix \mathbf{J}_F of transformation \mathbf{F} . Equation 6.4 shows \mathbf{J}_F as the product of the Jacobians of the transformations that make up \mathbf{F} .

$$\dot{\mathbf{x}}_i = \mathbf{J}_F \dot{\Upsilon} \quad (6.3)$$

$$\mathbf{J}_F = \mathbf{J}_{f_i} \cdot \mathbf{J}_P \cdot \mathbf{J}_{f_h} \cdot \mathbf{J}_T \quad (6.4)$$

Function \mathbf{T} represents a change of coordinates. It must be bijective which guarantees that it admits an inverse. Assume that Γ is the inverse function of \mathbf{T} ($\Gamma = \mathbf{T}^{-1}$). Function Γ , from \mathbf{R}^3 into \mathbf{R}^3 , transforms cartesian coordinates \mathbf{X}_w in new coordinates Υ (equation 6.5). \mathbf{J}_Γ is the jacobian matrix of Γ (equation 6.6). If \mathbf{T} is injective then the jacobian matrix \mathbf{J}_T is non-singular with inverse \mathbf{J}_T^{-1} [26]. Replacing \mathbf{J}_T by \mathbf{J}_Γ^{-1} in equation 6.4 yields equation 6.7 showing the jacobian matrix of \mathbf{F} expressed in terms of the scalar function of Γ and its partial derivatives.

$$\Gamma(\mathbf{X}_w) = (\phi(X, Y, Z), \psi(X, Y, Z), \rho(X, Y, Z))^t \quad (6.5)$$

$$\mathbf{J}_\Gamma = \begin{bmatrix} \phi_X & \phi_Y & \phi_Z \\ \psi_X & \psi_Y & \psi_Z \\ \rho_X & \rho_Y & \rho_Z \end{bmatrix} \quad (6.6)$$

$$\mathbf{J}_F = \mathbf{J}_{f_i} \cdot \mathbf{J}_P \cdot \mathbf{J}_{f_h} \cdot \mathbf{J}_\Gamma^{-1} \quad (6.7)$$

6.1.2 Criteria to Select the World Coordinate System

Function \mathbf{F} is a transformation from \mathbf{R}^3 (3D world space) into \mathbf{R}^2 (image plane). In equations 6.8 and 6.9 \mathbf{F} and \mathbf{J}_F are written in terms of scalar functions and their partial derivatives. The relationship between world and image points can be complex and counter intuitive. The mathematical expression of the mapping function \mathbf{F} depends on the transformation \mathbf{T} (see equations 6.2, 6.4 and 6.7). The selection

6. A General Framework for Selecting the World Coordinate System

of a certain coordinate system to reference points in the scene changes the way \mathbf{F} is written but does not change the intrinsic nature of the mapping. However an adequate choice of the world may lead to simpler mathematical relations between position and velocity in space and position and velocity in image. In this section we discuss criteria for the selection of the world coordinate system.

$$\mathbf{F}(\boldsymbol{\Upsilon}) = (h(\phi, \psi, \rho), g(\phi, \psi, \rho))^t \quad (6.8)$$

$$\mathbf{J}_{\mathbf{F}} = \begin{bmatrix} h_{\phi} & h_{\psi} & h_{\rho} \\ g_{\phi} & g_{\psi} & g_{\rho} \end{bmatrix} \quad (6.9)$$

The Compactness Constraint

Consider the general central projection as a mapping of 3D points, expressed in Cartesian coordinates $\mathbf{X}_w = (X, Y, Z)^t$, into 2D image points $\mathbf{x}_i = (x_i, y_i)$. The transformation is a function from \mathbf{R}^3 into \mathbf{R}^2 with loss of information (depth). In general the two coordinates in the image plane depend on the three coordinates in space. The image provides partial information about each one of the three world coordinates but we are not able to fully recover any of those parameters without further constraints. The imaging process implies loss of information and there is no additional transformation \mathbf{T} that can change that. However it would be advantageous that image coordinates depend only on two of the 3D parameters. In many situations that can be achieved by means of a change of coordinates \mathbf{T} . The change of world coordinates must be performed in such a way that \mathbf{F} only depends on two of those coordinates. Assuming that $\boldsymbol{\Upsilon} = (\phi, \psi, \rho)$ are the new 3D coordinates, \mathbf{F} becomes a function of only ϕ and ψ whenever the partial derivatives h_{ρ} and g_{ρ} are zero. If a certain change of coordinates \mathbf{T} leads to a jacobian matrix $\mathbf{J}_{\mathbf{F}}$ with a zero column then it is said that mapping \mathbf{F} is in a compact form and coordinate transformation \mathbf{T} verifies the "compactness constraint".

Assume that a world coordinate system satisfying the "compactness constraint" is selected. If equation 6.10 is verified then the image coordinates (x_i, y_i) depend only on (ϕ, ψ) and \mathbf{F} becomes a function from \mathbf{R}^2 in \mathbf{R}^2 ($\mathbf{x}_i = \mathbf{F}(\boldsymbol{\Upsilon}_c)$ with $\boldsymbol{\Upsilon}_c = (\phi, \psi)^t$). A function from \mathbf{R}^3 into \mathbf{R}^2 is never invertible, thus putting \mathbf{F}

6.1. World Coordinates for Static Imaging Systems

in a compact form is a necessary condition to find out an inverse mapping \mathbf{F}^{-1} . If \mathbf{F}^{-1} exists then two of the three 3D parameters of motion can be recovered from image ($\mathbf{r}_c = \mathbf{F}^{-1}(\mathbf{x}_i)$) and the jacobian matrix \mathbf{J}_F can be written in term of image coordinates x_i and y_i . By verifying the "compactness constraint" the relationships in position and velocity between the 3D world and the image plane tend to be more compact and intuitive and vision yields all the information about two of the 3D world coordinates and none about the third one.

$$h_\rho = 0 \quad \wedge \quad g_\rho = 0 \quad (6.10)$$

The Decoupling Constraint

Assume that the "compactness constraint" is verified. This means that a coordinate transformation \mathbf{T} is used such that image coordinates (x_i, y_i) depend only on (ϕ, ψ) . It would be also advantageous to define a world coordinate system such that x_i depends only of ϕ and y_i depends only of ψ . This is equivalent to say that h_ψ and g_ϕ are both zero. The one to one correspondence is an advantageous feature allowing a better understanding of the imaging process and simplifying subsequent calculations. If a coordinate transformation \mathbf{T} is used such that both equations 6.10 and 6.11 are verified then it is said that \mathbf{F} is in a compact and decoupled form and that \mathbf{T} verifies both the "compactness constraint" and the "decoupling constraint".

$$h_\psi = 0 \quad \wedge \quad g_\phi = 0 \quad (6.11)$$

In short, given a general central projection mapping, the goal is to select a coordinate transformation \mathbf{T} verifying both:

- the "compactness constraint" (equation 6.10)
- the "decoupling constraint" (equation 6.11)

The coordinate system used to reference points in the scene does not change the intrinsic nature of the mapping nor introduces any additional information. There are situations where it is impossible to find a world coordinates transformation that verifies the "compactness constraint" and/or the "decoupling constraint".

6. A General Framework for Selecting the World Coordinate System

Methodologies to find out if it exists such a transformation will be introduced latter.

6.2 Applying the Framework to Static Imaging Systems

The previous section derived a general framework to model central projection systems and select the most suitable coordinates to reference the world. This section applies the framework to the situations of a conventional perspective camera and central catadioptric system.

6.2.1 Conventional Perspective Camera

Consider image acquisition performed by a static conventional perspective camera. The image formation process follows the scheme depicted in Fig. 6.1 where function \mathbf{f}_i is given by equation 6.12. Assume that the matrix of intrinsic parameters is $\mathbf{K}_c = \mathbf{I}$ and $\mathbf{P} = [\mathbf{I}|\mathbf{0}]$ (the origin of the cartesian reference frame is coincident with the camera center and the image plane is perpendicular to the Z axis). This section derives a world coordinate system that verifies both the compactness and decoupling constraint. If nothing is stated we will work with the inverse transformation Γ instead of the direct transformation \mathbf{T} .

$$\mathbf{f}_i() : (x, y, z) \longrightarrow \left(\frac{x}{z}, \frac{y}{z}\right) \quad (6.12)$$

Constraining Γ to Obtain a New World Coordinate System

Functions \mathbf{f}_i , \mathbf{P} and \mathbf{f}_h , as well as their jacobian matrices, are defined for the perspective camera case. Replacing \mathbf{J}_Γ (equation 6.6) in equation 6.7 yields $\mathbf{J}_\mathbf{F}$ in terms of the partial derivatives of the scalar functions of Γ (the computation is omitted). If \mathbf{F} is in a compact form then the third column of $\mathbf{J}_\mathbf{F}$ must be zero (equation 6.10) which leads to equations 6.13. A transformation of coordinates Γ that verifies the compactness constraint can be computed by solving the partial differential equations 6.13 with respect to the scalar functions ϕ , ψ and ρ (equation 6.5).

6.2. Applying the Framework to Static Imaging Systems

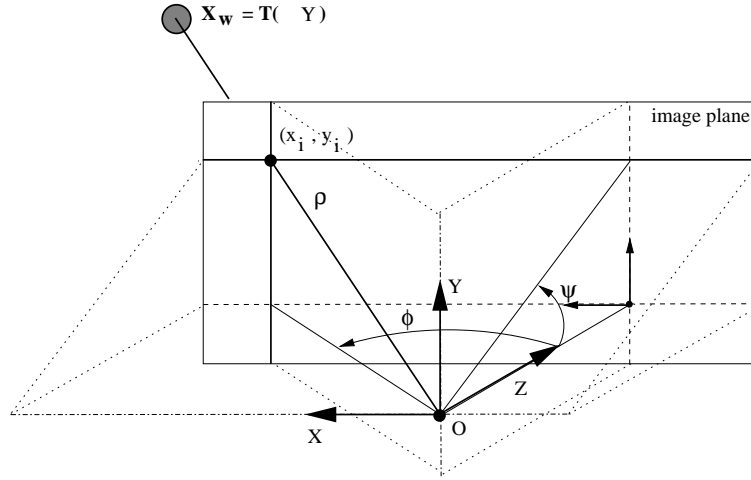


Figure 6.2: Conventional Perspective Camera Case

$$\begin{cases} Z(\phi_Y\psi_Z - \phi_Z\psi_Y) + X(\phi_Y\psi_X - \phi_X\psi_Y) = 0 \\ Z(\phi_Z\psi_X - \phi_X\psi_Z) + Y(\phi_Y\psi_X - \phi_X\psi_Y) = 0 \end{cases} \quad (6.13)$$

The partial differential equations corresponding to the "decoupling constraint" can be derived in a similar way. If the mapping F is decoupled then both h_ψ and g_ϕ must be zero, which leads to equation 6.14. A world coordinate transformation Γ verifying both the compactness and the decoupling constraint can be computed by solving simultaneously equations 6.13 and 6.14. Nevertheless the integration of systems of partial differential equations can be difficult and in general it generates many solutions. Adequate coordinate systems will be derived by geometrical means. Equations 6.13 and 6.14 will be used to prove that the selected coordinate transformation verifies the compactness and/or decoupling constraints.

$$\begin{cases} Z(\phi_Z\rho_Y - \phi_Y\rho_Z) + X(\phi_Y\rho_X - \phi_X\rho_Y) = 0 \\ Z(\psi_Z\rho_X - \psi_X\rho_Z) + Y(\psi_Y\rho_X - \psi_X\rho_Y) = 0 \end{cases} \quad (6.14)$$

Fig. 6.2 is a representation of the image formation process for the specified perspective camera situation. A point in the scene with cartesian coordinates (X, Y, Z) is projected in the image plane at position (x_i, y_i) . It is assumed that only the points in front of the camera are visible ($Z > 0$). The world points that can be projected in a certain image vertical line, lie in a vertical plane rotated

6. A General Framework for Selecting the World Coordinate System

around the camera referential Y axis, containing the center of projection. In a similar manner, the world points projected in an horizontal line of the image belong to an horizontal plane rotated around the X axis. Consider the pencil of the vertical planes where each one is indexed by the corresponding rotation angle ϕ . There is a one to one correspondence between the vertical lines in the image and the ϕ angles. In a similar manner, a pencil of horizontal planes, indexed by the corresponding rotation angle ψ , is defined. Each horizontal line in the image is associated to one ψ angle. Each pair (ϕ, ψ) defines a pair of planes that intersect in a projective ray. If the depth of the world point along the ray is ρ (always positive), the set of values (ϕ, ψ, ρ) gives the position of the world point in an unique way. We have derived a new system of coordinates to represent points in the 3D space. Equation 6.15 establishes the relationship between (ϕ, ψ, ρ) and the conventional Cartesian coordinates (X, Y, Z) .

$$\begin{cases} \phi = \arctan\left(\frac{X}{Z}\right) \\ \psi = \arctan\left(-\frac{Y}{Z}\right) \\ \rho = \sqrt{X^2 + Y^2 + Z^2} \end{cases} \quad (6.15)$$

Equation 6.16 gives the jacobian matrix of the derived coordinate transformation Γ . The proposed change of coordinates is a solution of the set of differential equations 6.13 and 6.14. Γ satisfies both the compactness and decoupling constraint for the static perspective camera case.

$$\mathbf{J}_\Gamma = \begin{bmatrix} \frac{Z}{X^2+Z^2} & 0 & -\frac{X}{X^2+Z^2} \\ 0 & -\frac{Z}{Y^2+Z^2} & \frac{Y}{Y^2+Z^2} \\ \frac{X}{\sqrt{X^2+Y^2+Z^2}} & \frac{Y}{\sqrt{X^2+Y^2+Z^2}} & \frac{Z}{\sqrt{X^2+Y^2+Z^2}} \end{bmatrix} \quad (6.16)$$

Table 6.1 summarizes the results obtained in this section. Γ is a world coordinate transformation verifying both the compactness and decoupling constraint. Notice that the new coordinate system is different from the well known spherical coordinates. \mathbf{T} is the inverse function of Γ . Replacing \mathbf{T} in equation 6.2 (\mathbf{f}_i , \mathbf{P} and \mathbf{f}_h were already defined) leads to the mathematical expression of global mapping \mathbf{F} using the new coordinates. The jacobian matrix $\mathbf{J}_\mathbf{F}$ is obtained replacing \mathbf{J}_Γ in equation 6.7 by the result of equation 6.16.

6.2. Applying the Framework to Static Imaging Systems

$$\begin{aligned}
 \Gamma(\mathbf{X}_w) &= \left(\arctan\left(\frac{X}{Z}\right), \arctan\left(-\frac{Y}{Z}\right), \sqrt{X^2 + Y^2 + Z^2} \right)^t \\
 \mathbf{T}(\Upsilon) &= \frac{\rho}{\sqrt{1+\tan(\phi)^2+\tan(\psi)^2}} (\tan(\phi), -\tan(\psi), 1)^t \\
 \mathbf{F}(\Upsilon) &= (\tan(\phi), -\tan(\psi))^t \\
 \mathbf{J}_F(\Upsilon) &= \begin{bmatrix} \frac{1}{\cos(\phi)^2} & 0 & 0 \\ 0 & -\frac{1}{\cos(\psi)^2} & 0 \end{bmatrix}
 \end{aligned}$$

Table 6.1: Using a new coordinate system for perspective camera

$$\begin{aligned}
 \mathbf{F}(\mathbf{X}_w) &= \left(\frac{X}{Z}, \frac{Y}{Z} \right)^t \\
 \mathbf{J}_F(\mathbf{X}_w) &= \begin{bmatrix} \frac{1}{Z} & 0 & -\frac{X}{Z^2} \\ 0 & \frac{1}{Z} & -\frac{Y}{Z^2} \end{bmatrix}
 \end{aligned}$$

Table 6.2: Using a cartesian coordinate system for perspective camera

Applications

The global mapping \mathbf{F} and its jacobian matrix establish the relationship between position/velocity in space and position/velocity in image for the conventional camera situation. Assume that both image position \mathbf{x}_i and velocity $\dot{\mathbf{x}}_i$ are known. It was stated that there is a loss of information in the image formation process. Therefore it is not possible to fully recover 3D target motion from images without further information. Nevertheless, using a world coordinate system verifying the compactness and decoupling constraint, it is possible to partially recover the 3D parameters of motion in a straightforward manner.

Table 6.1 shows the mapping \mathbf{F} and the corresponding jacobian matrix \mathbf{J}_F written in terms of derived system of coordinates. Table 6.2 gives the mathematical expressions of \mathbf{F} and \mathbf{J}_F using world cartesian coordinates. In the former situation \mathbf{F} is in a compact form and it is possible to invert the mapping to recover position information. One obtains that $\phi = \arctan(x_i)$ and $\psi = -\arctan(y_i)$. When using cartesian coordinates, \mathbf{F} appears as a function of \mathbf{R}^3 in \mathbf{R}^2 , and the inversion is not possible.

6. A General Framework for Selecting the World Coordinate System

Equation 6.17 shows \mathbf{J}_F written in terms of image coordinates. It is derived by replacing in the jacobian matrix of Table 6.1 (ϕ, ψ) by $(\arctan(x_i), -\arctan(y_i))$. Knowing both position and velocity in the image one obtains $\dot{\phi} = (1 + x_i^2)^{-1}\dot{x}_i$ and $\dot{\psi} = -(1 + y_i^2)^{-1}\dot{y}_i$. Using the derived world coordinate system it is possible to partially recover the position and velocity of the target in the scene.

$$\mathbf{J}_F(\mathbf{x}_i) = \begin{bmatrix} 1 + x_i^2 & 0 & 0 \\ 0 & -(1 + y_i^2) & 0 \end{bmatrix} \quad (6.17)$$

6.2.2 Central Catadioptric Imaging System

A catadioptric realization of omnidirectional vision combines reflective surfaces and lenses. Central catadioptric imaging can be highly advantageous for many applications because it combines two important features: a single projection center and a wide field of view. The drawback of this type of sensors is that in general the mapping between points in the 3D world and in the image is highly non-linear. We wish to study the advantages of working directly with the catadioptric images without warping them. The proposed framework is used to derive a general transformation of coordinates that leads to a mapping between points in the world and in the image that verifies both the compactness and the decoupling constraints. Some applications are presented and discussed.

General Model for Central Projection Systems

An unifying theory for central catadioptric image formation has been presented in chapter 2. This section reviews the mapping model schematized in Fig. 2.3 and introduces the assumptions and notation used in the remain of the chapter.

Consider a generic scene point, visible by the catadioptric system, with cartesian coordinates \mathbf{X}_w in the world reference frame. The corresponding homogeneous representation is \mathbf{X}_h . Visible points in the scene \mathbf{X}_h are mapped into projective rays/points \mathbf{x} in the catadioptric system reference frame centered in the effective viewpoint. The transformation is linear being described by a 3×4 matrix \mathbf{P} (if nothing is stated it is assumed that $\mathbf{P} = [\mathbf{I}|\mathbf{0}]$). To each oriented projective ray/point $\mathbf{x} = (x, y, z)^t$, corresponds a projective ray/point $\hat{\mathbf{x}} = (\hat{x}, \hat{y}, \hat{z})^t$ in a coordinate system whose origin is in the camera projection center(Fig. 2.3). Assume

6.2. Applying the Framework to Static Imaging Systems

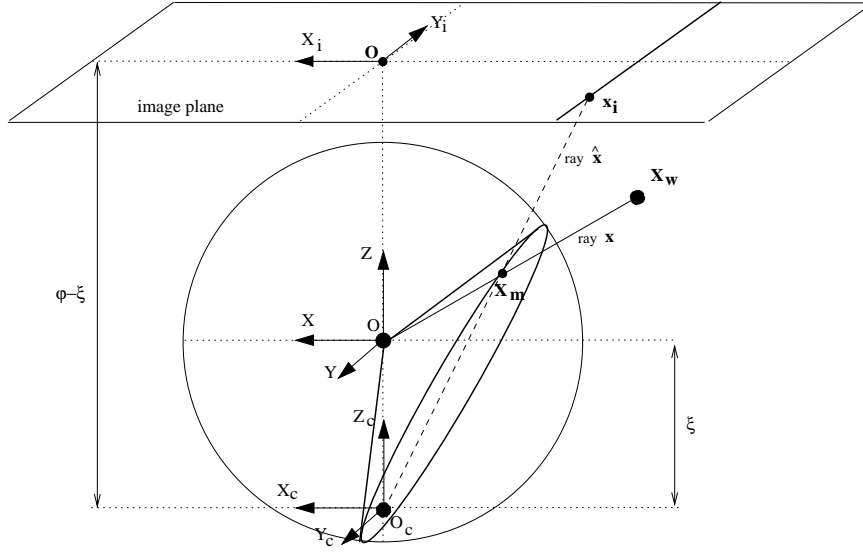


Figure 6.3: The sphere model for general central catadioptric image projection

that the matrix of intrinsic parameters is $\mathbf{K}_c = \mathbf{I}$ and that the camera is not rotated with relation to the reflective surface ($\mathbf{R}_c = \mathbf{I}$). From equations 2.4, 2.5 and 2.6 yields

$$\hat{\mathbf{x}} = \left(\frac{(\varphi - \xi)x}{\sqrt{x^2 + y^2 + z^2}}, -\frac{(\varphi - \xi)y}{\sqrt{x^2 + y^2 + z^2}}, \frac{z}{\sqrt{x^2 + y^2 + z^2}} + \xi \right)^t$$

Since the cartesian coordinates of the point in the catadioptric image plane are $\mathbf{x}_i = \left(\frac{\hat{x}}{\hat{z}}, \frac{\hat{y}}{\hat{z}} \right)^t$ then the function \mathbf{f}_i , referenced in the scheme of Fig. 6.1, is provided by equation 6.19. Function \mathbf{f}_i depends on mirror parameters ξ and φ . Table 2.3 shows these parameters for the different situations of central catadioptric imaging.

$$\mathbf{f}_i() : (x, y, z) \longrightarrow \left(\frac{(\varphi - \xi)x}{z + \xi\sqrt{x^2 + y^2 + z^2}}, -\frac{(\varphi - \xi)y}{z + \xi\sqrt{x^2 + y^2 + z^2}} \right) \quad (6.19)$$

Fig. 6.3 depicts the intuitive “concrete” model. To each visible point in space corresponds an oriented projective ray \mathbf{x} joining the 3D point with the effective projection center \mathbf{O} . The projective ray intersects a unit sphere centered in \mathbf{O} in a unique point \mathbf{X}_m . Consider a point \mathbf{O}_c with coordinates $(0, 0, -\xi)^t$ in sphere reference frame \mathfrak{R} . To each \mathbf{x} corresponds an oriented projective ray $\hat{\mathbf{x}}$ joining \mathbf{O}_c

6. A General Framework for Selecting the World Coordinate System

with the intersection point \mathbf{X}_m in the sphere surface. Assume that the catadioptric image plane is the horizontal plane $Z = \varphi - 2\xi$. The projective ray $\hat{\mathbf{x}}$ intersects the plane at \mathbf{x}_i which are the coordinates of the image point. The scene is projected into the sphere surface and then points on the sphere are re-projected into the catadioptric image plane from a novel projection center \mathbf{O}_c . Point $\mathbf{O}_c = (0, 0, -\xi)^t$ only depends on the mirror parameters (see Table.2.3). Notice that if $\xi = 0$ then \mathbf{O}_c is coincident with \mathbf{O} and function \mathbf{f}_i becomes similar to the one provided in equation 6.12.

The New World Coordinate System

The image formation process of a general central catadioptric system fits the scheme depicted in Fig. 6.1. The difference from the previous case of the perspective camera is that function \mathbf{f}_i is given by equation 6.19 instead of equation 6.12 where ξ and φ depend on the mirror parameters (see Table 2.3). The goal of this section is to derive a coordinate transformation Γ for which the global mapping \mathbf{F} between points in the world and in the catadioptric image is in a compact and decoupled form. Differential constraints similar to equations 6.13 and 6.14 can be derived in the same manner for this more general situation. As already mentioned, integration of partial differential equations can be a complex task, leading to multiple solutions. Once again the suitable new coordinates are derived geometrically and the differential constraints are used to confirm the results.

In Fig. 6.3 consider the vertical line in the catadioptric image plane parallel to the Y axis. All the points in the world that can be projected in this line lie in a conic surface, with the vertex coincident with the effective viewpoint \mathbf{O} . The conic surface intersects the unit sphere on a circumference, passing through the projection point \mathbf{X}_m , which limits a circle containing the point $\mathbf{O}_c = (0, 0, -\xi)^t$. The axis of the conic surface is always contained in the XOZ plane. In a similar way, the world points that can be projected in an horizontal line in the image lie in a conic surface. The difference is that the axis is now in the YOZ plane. We will call the first pencil of conic surfaces (with the axis in XOZ) the vertical pencil and the second (with the axis in YOZ) the horizontal pencil. A vertical and an horizontal conic surface intersect in two oriented projection rays. This ambiguity can be solved assuming that the camera used to acquire the catadioptric

6.2. Applying the Framework to Static Imaging Systems

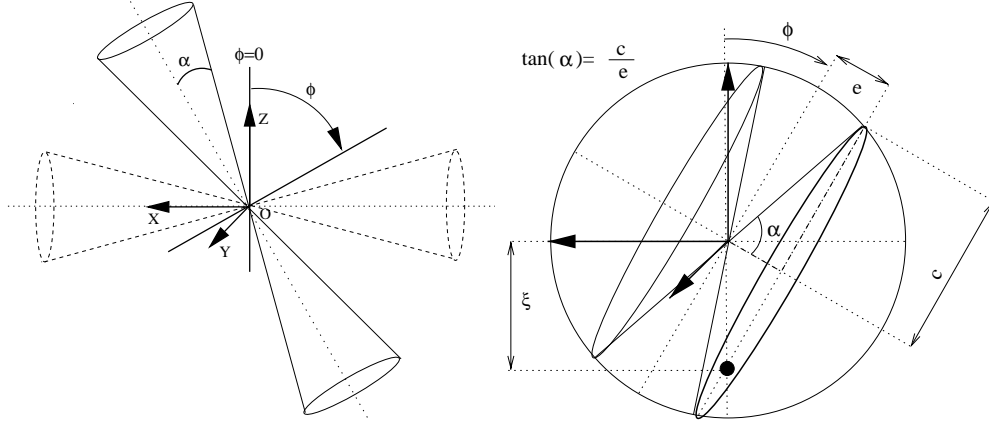


Figure 6.4: Left: Vertical and horizontal pencils of conic surfaces. Intersection of the lower and upper cone with the unit sphere

images is forward looking and that only points with $Z > -\xi$ are visible in the image plane. This section defines a coordinate system based in these two pencils of cones. It is shown that the resulting transformation of coordinates Γ verifies both the compactness and decoupling constraint for the general central projection situation.

Consider the dotted conic surface depicted in Fig. 6.4. Its axis coincides with the referential X axis, the vertex coincides with the origin and the aperture angle is α . World points (X, Y, Z) lying on the conic surface verify equation 6.20. Consider now the conic surface rotated around the Y axis by an angle ϕ . Equation 6.21 is obtained multiplying point coordinates by rotation matrix $e^{\phi \hat{y}}$ (where \hat{y} is the unit vector of the Y axis) [51] and replacing them in equation 6.20. Equation 6.21 defines the rotated conic surface depicted in Fig 6.4(Left). Angle α is the aperture angle and ϕ is the rotation angle around Y axis that can be used to index the cones in the vertical pencil described above.

$$Z^2 + Y^2 = X^2 \tan(\alpha)^2 \quad (6.20)$$

$$(X \sin(\phi) + Z \cos(\phi))^2 - (X \cos(\phi) - Z \sin(\phi))^2 \tan(\alpha)^2 + Y^2 = 0 \quad (6.21)$$

6. A General Framework for Selecting the World Coordinate System

Observe Fig. 6.4(Right) where the rotated conic surface of equation 6.21, the unit sphere and the re-projection center O_c are depicted. The lower cone must intersect the unit sphere on a circle containing the point $(0, 0, -\xi)$. World scene points lying in this cone project on a vertical line at the catadioptric image plane (see Fig. 6.3). Fig. 6.4(Right) shows that the aperture angle of the conic surface must be $\tan(\alpha) = \frac{c}{e}$ where $e^2 = \xi^2 \sin^2(\phi)$ and $c^2 = 1 - e^2$ (an unit sphere is assumed). Equation 6.22 is obtained replacing this result in equation 6.21.

$$\xi^2 \sin(\phi)^2 (X^2 + Y^2 + Z^2) - (X \cos(\phi) - Z \sin(\phi))^2 = 0 \quad (6.22)$$

Solving equation 6.22 with respect to ϕ we are able to compute the vertical conic surface with the desired features that contains a certain world point (X, Y, Z) . Notice however that it is a second order equation, thus for each point in 3D space, there are two ϕ solutions. Each solution defines a conic surface containing the point and intersecting the unit sphere in a circle passing through its projection X_m at the sphere surface. Nevertheless one of these circles contains point $(0, 0, \xi)^t$, while the other contains point $(0, 0, -\xi)^t$. The second solution is the one that must be used.

The ψ angle is used in the same manner to represent the horizontal pencil of conic surfaces. The derivation of the relationship between ψ and point world coordinates (X, Y, Z) is similar.

At this point, given a world point represented in Cartesian coordinates (X, Y, Z) we are able to compute a vertical and an horizontal conic surface, respectively referenced in an unique way by ϕ and ψ , that contain the point. In general the two cones intersect in two projective rays. However only one of them is visible in the image. The point lies in this projective ray, and a third coordinate ρ , that is the distance from the center, must be introduced. Equation 6.23 yields the derived change of coordinates Γ . Both ϕ and ψ are in the range $[-\pi/2, \pi/2]$.

$$\begin{cases} \phi = \arctan\left(\frac{X}{Z + \xi \sqrt{X^2 + Y^2 + Z^2}}\right) \\ \psi = \arctan\left(-\frac{Y}{Z + \xi \sqrt{X^2 + Y^2 + Z^2}}\right) \\ \rho = \sqrt{X^2 + Y^2 + Z^2} \end{cases} \quad (6.23)$$

Table 6.3 summarizes the results obtained so far for general central projection

6.2. Applying the Framework to Static Imaging Systems

$\mathbf{\Gamma}(\mathbf{X}_w)$	=	$\begin{bmatrix} \arctan\left(\frac{X}{Z+\xi\sqrt{X^2+Y^2+Z^2}}\right) \\ \arctan\left(-\frac{Y}{Z+\xi\sqrt{X^2+Y^2+Z^2}}\right) \\ \sqrt{X^2+Y^2+Z^2} \end{bmatrix}$
$\mathbf{T}(\mathbf{\Upsilon})$	=	$\begin{bmatrix} \rho \frac{l+\sqrt{1+(1-\xi^2)(\tan(\phi)^2+\tan(\psi)^2)}}{1+\tan(\phi)^2+\tan(\psi)^2} \tan(\phi) \\ -\rho \frac{l+\sqrt{1+(1-\xi^2)(\tan(\phi)^2+\tan(\psi)^2)}}{1+\tan(\phi)^2+\tan(\psi)^2} \tan(\psi) \\ \rho \frac{\sqrt{1+(1-\xi^2)(\tan(\phi)^2+\tan(\psi)^2)}-\xi(\tan(\phi)^2+\tan(\psi)^2)}{1+\tan(\phi)^2+\tan(\psi)^2} \end{bmatrix}$
$\mathbf{F}(\mathbf{\Upsilon})$	=	$(\varphi - \xi) \begin{bmatrix} \tan(\phi) \\ \tan(\psi) \end{bmatrix}$
$\mathbf{J}_F(\mathbf{\Upsilon})$	=	$(\varphi - \xi) \begin{bmatrix} \frac{1}{\cos(\phi)^2} & 0 & 0 \\ 0 & \frac{1}{\cos(\psi)^2} & 0 \end{bmatrix}$

Table 6.3: Using a new coordinate system for general central catadioptric imaging

systems. $\mathbf{\Gamma}$ is the derived world coordinate transformation that maps cartesian coordinates \mathbf{X}_w in the new coordinates $\mathbf{\Upsilon}$ associated with the conic surfaces reference frame. \mathbf{T} is the inverse function of $\mathbf{\Gamma}$. Replacing \mathbf{T} in equation 6.2 (\mathbf{f}_i , \mathbf{P} and \mathbf{f}_h were already defined) yields the mathematical expression of global mapping \mathbf{F} using the new coordinates. The corresponding jacobian matrix is \mathbf{J}_F . Accordingly equations 6.10 and 6.11, the derived coordinate transformation $\mathbf{\Gamma}$ verify both the compactness and decoupling constraints. Notice that if $\xi = 0$ then we have the perspective camera situation studied previously. The aperture of the reference conic surfaces is always $\pi/2$ and the change of coordinates 6.23 becomes equal to 6.15. The pencils of conic surfaces degenerate in the pencils of planes previously derived. The coordinate transformation proposed for the perspective camera situation is a particular case of this more general solution.

Applications

Applying the derived coordinate transformation $\mathbf{\Gamma}$ does not modify the image formation process nor introduces new information in the problem. Instead the geometric transformations are represented in a way that enables newer insights in the imaging process becoming more suitable to develop certain applications. This section shows some advantages in expressing the global mapping function in terms of an adequate system of world coordinates.

6. A General Framework for Selecting the World Coordinate System

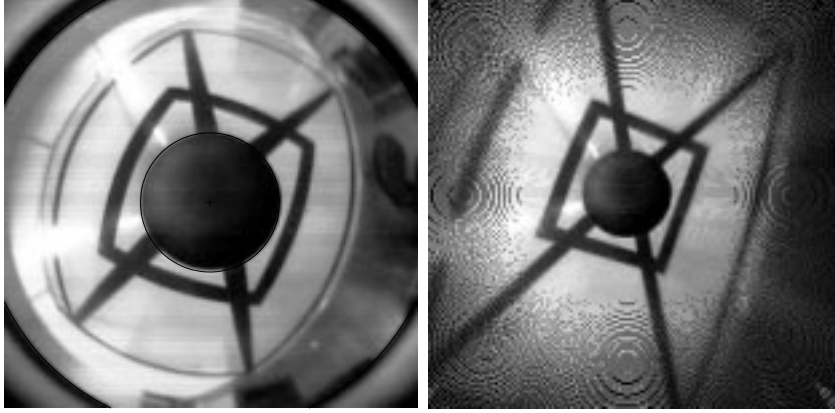


Figure 6.5: Generating a geometrically correct perspective image.

Using the derived world coordinates Υ , the global mapping \mathbf{F} is written in a compact and decoupled form (see table 6.3). Knowing both target image position \mathbf{x}_i and velocity $\dot{\mathbf{x}}_i$ the recovery of 3D parameters of motion is straightforward. As a result $\phi = \arctan(x_i)$ and $\psi = \arctan(y_i)$. Replacing these parameters in the jacobian matrix of table 6.3 yields the jacobian matrix of the global mapping \mathbf{F} written in terms of image coordinates (equation 6.24). It follows that $\dot{\phi} = (1 + \frac{x_i^2}{(\varphi-\xi)^2})^{-1} \frac{\dot{x}_i}{\varphi-\xi}$ and $\dot{\psi} = -(1 + \frac{y_i^2}{(\varphi-\xi)^2})^{-1} \frac{\dot{y}_i}{\varphi-\xi}$. Similarly to the perspective camera case, catadioptric images do not provide any information about the third parameter of motion ρ .

$$\mathbf{J}_{\mathbf{F}}(\mathbf{x}_i) = (\varphi - \xi) \begin{bmatrix} 1 + \frac{x_i^2}{(\varphi-\xi)^2} & 0 & 0 \\ 0 & 1 + \frac{y_i^2}{(\varphi-\xi)^2} & 0 \end{bmatrix} \quad (6.24)$$

It is a well known result that it is always possible to generate a geometrically correct perspective image from an image acquired by a central projection catadioptric system[5]. The first image of Fig. 6.5 was acquired by an omnidirectional system that combines a parabolic mirror with an orthographic camera. A generic scene point with cartesian coordinates \mathbf{X}_w is projected at position \mathbf{x}_i in the catadioptric image plane (see Fig.6.3). The field of view is nearly 180° , thus all points in the world such that $Z > 0$ are visible. Assume a conventional perspective camera with projection center at the effective viewpoint $\mathbf{0}$ and optical axis aligned with the Z-axis. A generic world point with coordinates $\mathbf{X}_w = (X, Y, Z)^t$

6.2. Applying the Framework to Static Imaging Systems

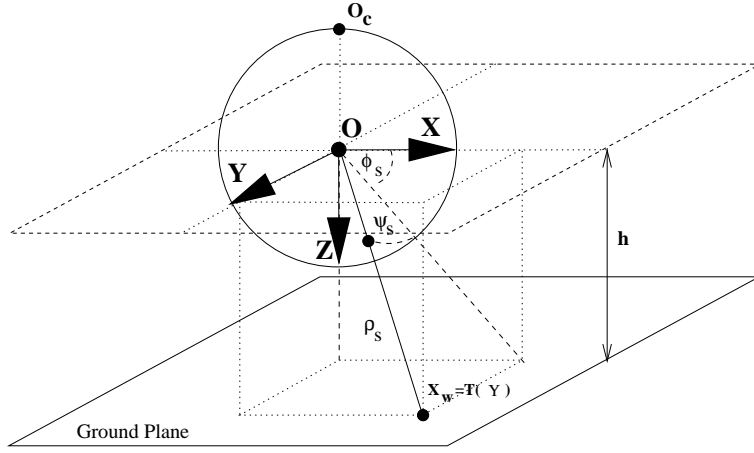


Figure 6.6: Surveillance application using central panoramic imaging.

is projected in the perspective image at position $\mathbf{x}_p = (\frac{X}{Z}, \frac{Y}{Z})^t$. The goal is to derive the function \mathbf{f}_p that maps points in the catadioptric image plane into points in the perspective image plane ($\mathbf{x}_p = \mathbf{f}_p(\mathbf{x}_i)$). This mapping function can be determined in a straightforward manner using the results of table 6.3. As already mentioned $\phi = \arctan(x_i)$ and $\psi = \arctan(y_i)$. Replacing ϕ and ψ in the coordinate transformation \mathbf{T} we are able to compute world coordinates \mathbf{X}_w as a function of catadioptric image coordinates \mathbf{x}_i and ρ . Making $\mathbf{x}_p = (\frac{X}{Z}, \frac{Y}{Z})^t$ the dependence on ρ disappears and \mathbf{x}_p is obtained as a function of \mathbf{x}_i . Function \mathbf{f}_p is presented in equation 6.25 for the parabolic system case ($\xi = 1$). The geometrically correct perspective image generated from the derived mapping function is presented at Fig. 6.5.

$$\mathbf{f}_p() : (x_i, y_i) \longrightarrow \left(\frac{4px_i}{4p^2 - (x_i^2 + y_i^2)}, -\frac{4py_i}{4p^2 - (x_i^2 + y_i^2)} \right) \quad (6.25)$$

Spherical coordinates are broadly used in computer vision and robotics. For some applications it can be useful to reference world points using spherical coordinates. Assume, as an example, that we intend to use our omnidirectional camera in a surveillance application to obtain position information of a set of targets. The imaging system is fixed at the ceiling as depicted in Fig. 6.6. The target position can be referenced in a simple way using spherical coordinates $\Upsilon_s = (\phi_s, \psi_s, \rho_s)^t$. The goal is to derive the function \mathbf{f}_s which transforms catadioptric image coordinates

6. A General Framework for Selecting the World Coordinate System

ordinates \mathbf{x}_i , in the two world spherical coordinates (ϕ_s, ψ_s) that can be fully recovered $((\phi_s, \psi_s) = \mathbf{f}_s(\mathbf{x}_i))$ without further restrictions. $\mathbf{X}_w = (X, Y, Z)^t$ can be computed in terms of \mathbf{x}_i and ρ . Since $\phi = \arctan(x_i)$ and $\psi = \arctan(y_i)$, then $\mathbf{X}_w = \mathbf{T}(\arctan(x_i), \arctan(y_i), \rho)$ (see table 6.3). Replacing (X, Y, Z) in $\phi_s = \arctan(\frac{X}{Z})$ and $\psi_s = \arctan(\frac{Y}{\sqrt{X^2+Z^2}})$ by the obtained result, the dependence on ρ is eliminated and function \mathbf{f}_s is obtained (equation 6.26). In addition, if the height h of the room is known and if the imaged point is on the floor one obtains $\rho_s = h\sqrt{(4p^2 + x_i^2 + y_i^2)^2 / (4p^2 - x_i^2 - y_i^2)^2}$.

$$\mathbf{f}_s() : (x_i, y_i) \rightarrow \left(\arctan\left(\frac{4px_i}{4p^2 - (x_i^2 + y_i^2)}\right), \arctan\left(\frac{-4py_i}{\sqrt{(4p^2 + x_i^2 - y_i^2)^2 - 4x_i^2 y_i^2}}\right) \right) \quad (6.26)$$

6.3 World Coordinates for Imaging Systems with Motion

The two previous sections focuses on static central projection imaging systems. The global mapping \mathbf{F} from points in the 3D world in points in the 2D image plane is derived. It is shown that the mathematical expression of \mathbf{F} depends on the coordinate system used to reference points in the scene and that an adequate choice of coordinates presents several advantages. In this situation the imaging system does not move, thus motion in the image plane only depends on 3D motion in the scene.

This section focuses on active tracking of moving targets where the goal is to use visual information to control camera motion such that target image is kept constant. The mathematical framework derived in section 6.1 is extended to central projection imaging systems with rigid motion. Global mapping between 3D world points and 2D image points is derived. Criteria to select adequate world coordinates are discussed. The results obtained are applied to active visual tracking using three different platforms and/or imaging sensors.

As mentioned the assumed platforms have less than 3 DOF. To control such a constrained motion it is enough to work with 3D target position and velocity. If nothing is stated it is assumed that \mathbf{X}_w is a vector with target mass center position

6.3. World Coordinates for Imaging Systems with Motion

in cartesian coordinates, Υ represents the target position in the established alternative system of coordinates and \mathbf{x}_i is the vector with the target image coordinates. Vector Θ is introduced to represent camera/platform position in an inertial world reference frame.

6.3.1 Mapping points from the 3D world into the 2D image plane

Consider the schematic of Fig. 6.1 which depicts the general mapping \mathbf{F} performed by a static central projection imaging system. Consider $\mathbf{P} = \mathbf{R}[\mathbf{I} - \mathbf{C}]$, with \mathbf{R} the rotation matrix between camera and world and \mathbf{C} the projection center coordinates. Matrix \mathbf{P} depends on the pose of the imaging system in world coordinates. If the imaging system is static then the matrix \mathbf{P} is constant. On the other hand if the camera moves, its pose changes, and matrix \mathbf{P} is no longer constant. Assume the imaging system is mounted on a moving platform. The pose of the camera depends on the position Θ of the platform. Since matrix \mathbf{P} depends on the pose of the camera then \mathbf{P} is a function of Θ .

The mapping between the scene and the image plane depends on the target 3D coordinates Υ and on the camera pose parameterized by Θ . The target image coordinates \mathbf{x}_i are given by equation 6.27 where transformation \mathbf{F} can be written as the composition of equation 6.28. The difference between equations 6.2 and 6.28 is that matrix \mathbf{P} is no longer constant and appears as a function of Θ .

$$\mathbf{x}_i = \mathbf{F}(\Upsilon, \Theta) \quad (6.27)$$

$$\mathbf{F}(\Upsilon, \Theta) = \mathbf{f}_i(\mathbf{P}(\Theta) \cdot \mathbf{f}_h(\mathbf{T}(\Upsilon))) \quad (6.28)$$

The target velocity in the image $\dot{\mathbf{x}}_i$ is computed in 6.29. Equation 6.29 is obtained by differentiating equation 6.27 with respect to time. $\mathbf{J}_F = [\mathbf{J}_F^\Upsilon | \mathbf{J}_F^\Theta]$ is the jacobian matrix of function \mathbf{F} , $\dot{\Upsilon}$ is the target 3D velocity, and $\dot{\Theta}$ represents the camera/platform velocity. \mathbf{J}_F^Υ is given in 6.30 with \mathbf{J}_Γ the jacobian matrix of the inverse coordinate transformation Γ . \mathbf{J}_F^Θ is computed in 6.31 and does not depend on the jacobian matrix of the world coordinate transformation \mathbf{J}_Γ . The image velocity $\dot{\mathbf{x}}_i$ depends both on target velocity $\dot{\Upsilon}$ and on the camera/platform velocity $\dot{\Theta}$. The second term in equation 6.29 is known in the literature by egomotion and

6. A General Framework for Selecting the World Coordinate System

represents the image motion induced by camera/platform motion.

$$\dot{\mathbf{x}}_i = \mathbf{J}_F^Y \dot{\mathbf{Y}} + \mathbf{J}_F^\Theta \dot{\Theta} \quad (6.29)$$

$$\mathbf{J}_F^Y = \mathbf{J}_{f_i} \cdot \mathbf{J}_P^{X_h} \cdot \mathbf{J}_{f_h} \cdot \mathbf{J}_\Gamma^{-1} \quad (6.30)$$

$$\mathbf{J}_F^\Theta = \mathbf{J}_{f_i} \cdot \mathbf{J}_P^\Theta. \quad (6.31)$$

The target image depends on the relative position between camera and target. Describing target motion in a coordinate frame attached to the imaging device simplifies the position/velocity relationships of equations 6.27 to 6.31. The egomotion term disappears and the position/velocity in the image depends only on the position/velocity of the target in the camera system of coordinates. The platform position, provided, for example, by encoder readings, and joint control commands are usually defined in an inertial base coordinate frame. If the control input is defined in the camera coordinate system then the transformations between the two reference systems can not be avoided. In the position/velocity relationships of equations 6.27 to 6.31, both camera and target motion are described in a common inertial system of coordinates. Errors in the image plane can be directly related with the control inputs commands in the task space, thus there is no need of additional coordinate transformations. Multiple cameras can be integrated in a natural way by describing the target motion in a common coordinate frame. The explicit computation of an egomotion term can be used for image segmentation [60, 17, 15].

6.3.2 Criteria to Select the World Coordinate System

The mathematical expression of the global mapping \mathbf{F} depends on system of coordinates used to reference target position in the scene (equation 6.28). The intrinsic nature of the mapping does not depend on the selected coordinate frame. However, as seen in section 6.1 for static central catadioptric imaging systems, an adequate choice of the coordinate system can be highly advantageous. Criteria to select suitable coordinate transformations Γ for tracking applications are discussed in the present section.

6.3. World Coordinates for Imaging Systems with Motion

Analytical Solution for Visual Control of Motion in Tracking Applications

Consider the imaging system mounted on a moving platform. If nothing is stated assume a platform with 2 DOF (Θ is a 2×1 vector). To perform active tracking of a moving target, the platform motion must be controlled such that the target projection is kept in constant position/velocity in the image plane. Assume that \mathbf{x}_d and $\dot{\mathbf{x}}_d$ are the desired target position and velocity in the image. Consider that Θ_c and $\dot{\Theta}_c$ are the position and velocity commands sent to the platform actuators which have an unitary transfer function. Thus, Θ_c and $\dot{\Theta}_c$ must verify both $\mathbf{x}_d = \mathbf{F}(\Upsilon, \Theta_c)$ and $\dot{\mathbf{x}}_d = \mathbf{J}_F^\Upsilon \cdot \dot{\Upsilon} + \mathbf{J}_F^\Theta \cdot \dot{\Theta}_c$ with Υ and $\dot{\Upsilon}$ target 3D position and velocity (see equations 6.27 to 6.31).

If nothing is stated assume in the sequel that at each frame time instant both target position \mathbf{x}_i and velocity $\dot{\mathbf{x}}_i$ are measured in the image plane and that the platform position Θ and velocity $\dot{\Theta}$ are estimated using the encoder readings. The goal is to determine the position and velocity commands Θ_c and $\dot{\Theta}_c$ knowing \mathbf{x}_i , $\dot{\mathbf{x}}_i$, Θ and $\dot{\Theta}$.

Proposition 6.1: *Assume that it is possible to compute target 3D position Υ from target position in image \mathbf{x}_i and camera pose Θ . In a similar way consider that camera position Θ can be uniquely calculated given target position in image \mathbf{x}_i and in the scene Υ . If these two conditions hold then, given target position in image \mathbf{x}_i and camera pose Θ , it is possible to compute camera position Θ_c such that target is projected in a pre-defined position \mathbf{x}_d in the image plane. A similar statement can be made for velocity relationships.*

Proof: Function \mathbf{F} computes target position in image \mathbf{x}_i given target 3D position Υ and camera pose Θ (equation 6.27). Assume that it exists a function v which enables the computation of Υ given both \mathbf{x}_i and Θ (equation 6.32). In a similar way, using function θ , Θ can be calculated knowing both \mathbf{x}_i and Υ (equation 6.33).

$$\Upsilon = v(\mathbf{x}_i, \Theta) \quad (6.32)$$

$$\Theta = \theta(\mathbf{x}_i, \Upsilon) \quad (6.33)$$

6. A General Framework for Selecting the World Coordinate System

Assume that the target image position \mathbf{x}_i is measured and camera pose Θ is estimated from encoder readings. The world target 3D position is $v(\mathbf{x}_i, \Theta)$. The camera position Θ_c such that target image is projected in position \mathbf{x}_d is given in equation 6.34.

$$\Theta_c = \theta(\mathbf{x}_d, v(\mathbf{x}_i, \Theta)) \quad (6.34)$$

Equations 6.35 and 6.36 are obtained differentiating 6.32 and 6.33 with respect to time. $\mathbf{J}_v = [\mathbf{J}_v^i | \mathbf{J}_v^\Theta]$ and $\mathbf{J}_\theta = [\mathbf{J}_\theta^i | \mathbf{J}_\theta^\Theta]$ are the jacobian matrices of functions v and θ .

$$\dot{\Upsilon} = \mathbf{J}_v^i \dot{\mathbf{x}}_i + \mathbf{J}_v^\Theta \dot{\Theta} \quad (6.35)$$

$$\dot{\Theta} = \mathbf{J}_\theta^i \dot{\mathbf{x}}_i + \mathbf{J}_\theta^\Upsilon \dot{\Upsilon} \quad (6.36)$$

Knowing both the target velocity in the image $\dot{\mathbf{x}}_i$ and the camera velocity $\dot{\Theta}$ at the frame acquisition time instant, it is possible to compute the camera velocity $\dot{\Theta}_c$ such that target is projected in the image plane with desired velocity $\dot{\mathbf{x}}_d$. This results is shown in equation 6.37 derived by differentiating 6.34 with respect to time.

$$\dot{\Theta}_c = \mathbf{J}_\theta^i \dot{\mathbf{x}}_d + \mathbf{J}_\theta^\Upsilon \mathbf{J}_v^i \dot{\mathbf{x}}_i + \mathbf{J}_\theta^\Upsilon \mathbf{J}_v^\Theta \dot{\Theta} \quad (6.37)$$

Proposition 6.2: *Consider the global mapping function \mathbf{F} and the corresponding jacobian matrix $\mathbf{J}_\mathbf{F} = [\mathbf{J}_\mathbf{F}^\Upsilon | \mathbf{J}_\mathbf{F}^\Theta]$. If it exists a function v that enables the computation of the target 3D position from target image and camera pose, then $\mathbf{J}_\mathbf{F}^\Upsilon$ must be invertible. In a similar way if it exists a function θ that allows the calculation of the camera position from the target position in the image and in the world, then matrix $\mathbf{J}_\mathbf{F}^\Theta$ must have inverse.*

Proof: Assume that the function v exists (equation 6.32). The mathematical relationship of equation 6.38 is obtained replacing $\dot{\Upsilon}$ in equation 6.29 by the result of equation 6.35. Equation 6.39 is derived replacing in 6.35 $\dot{\mathbf{x}}_i$ by the result presented in 6.29.

6.3. World Coordinates for Imaging Systems with Motion

$$\dot{\mathbf{x}}_i = \mathbf{J}_F^{\Upsilon} \mathbf{J}_v^i \dot{\mathbf{x}}_i + (\mathbf{J}_F^{\Upsilon} \mathbf{J}_v^{\Theta} + \mathbf{J}_F^{\Theta}) \dot{\Theta} \quad (6.38)$$

$$\dot{\Upsilon} = \mathbf{J}_v^i \mathbf{J}_F^{\Upsilon} \dot{\Upsilon} + (\mathbf{J}_v^i \mathbf{J}_F^{\Theta} + \mathbf{J}_v^{\theta}) \dot{\Theta} \quad (6.39)$$

If function v exists then equalities 6.38 and 6.39 must hold. From these equalities come $\mathbf{J}_v^i \mathbf{J}_F^{\Upsilon} = \mathbf{J}_F^{\Upsilon} \mathbf{J}_v^i = \mathbf{I}$. Thus matrix \mathbf{J}_F^{Υ} must have inverse which is \mathbf{J}_v^i . Considering that $\mathbf{J}_F^{\Upsilon} \mathbf{J}_v^{\Theta} + \mathbf{J}_F^{\Theta} = \mathbf{J}_v^i \mathbf{J}_F^{\Theta} + \mathbf{J}_v^{\theta} = 0$ and making $\mathbf{J}_v^i = (\mathbf{J}_F^{\Upsilon})^{-1}$ comes the results of equation 6.40.

$$\begin{cases} \mathbf{J}_v^i = (\mathbf{J}_F^{\Upsilon})^{-1} \\ \mathbf{J}_v^{\Theta} = -(\mathbf{J}_F^{\Upsilon})^{-1} \mathbf{J}_F^{\Theta} \end{cases} \quad (6.40)$$

Consider now that function θ exists. In a similar way it can be proved that \mathbf{J}_F^{Θ} must be invertible and that \mathbf{J}_θ^i and $\mathbf{J}_\theta^{\Upsilon}$ are given by equation 6.41.

$$\begin{cases} \mathbf{J}_\theta^i = (\mathbf{J}_F^{\Theta})^{-1} \\ \mathbf{J}_\theta^{\Upsilon} = -(\mathbf{J}_F^{\Theta})^{-1} \mathbf{J}_F^{\Upsilon} \end{cases} \quad (6.41)$$

Equation 6.37 is rewritten in 6.42 using the results of equations 6.40 and 6.41

$$\dot{\Theta}_c = \dot{\Theta} + (\mathbf{J}_F^{\Theta})^{-1} (\dot{\mathbf{x}}_d - \dot{\mathbf{x}}_i) \quad (6.42)$$

Coordinate Transformation Constraints

In the previous section proposition 6.2 establishes the necessary conditions for the existence of functions v and θ and proposition 6.1 establishes the sufficient conditions for the existence of an analytical solution for the tracking problem.

Assuming a platform with 2 DOF, then \mathbf{J}_F^{Θ} is a 2×2 matrix. This matrix is invertible or not depending on the features of the system, in particular the kinematics of the robotic platform and the type of image sensor. The coordinate frame where the platform position Θ is defined is considered as a problem specification that can not be changed.

In general, the world target position is referenced by $\Upsilon = (\phi, \psi, \rho)^t$, which is a 3×1 vector. The corresponding jacobian matrix \mathbf{J}_F^{Υ} is a 2×3 matrix similar to the one shown in equation 6.9. A 2×3 matrix is never invertible and, accordingly to proposition 6.2, it is not possible to define a function v to recover target 3D

6. A General Framework for Selecting the World Coordinate System

position information. As mentioned in section 6.1, the image formation process transforms points in \mathbf{R}^3 into points in \mathbf{R}^2 . There is a loss of information and world target position can not be fully recovered from a single image. Nevertheless, as mentioned in 6.1.2, if \mathbf{F} is written in a compact form then Υ can be partially recovered in a straightforward manner. If the recovered target 3D parameters are enough to derive function θ then an analytical solution for the tracking problem can be reached (proposition 6.1).

If it exists a transformation of coordinates Γ verifying the compactness constraint then it is possible to write the global mapping \mathbf{F} in a compact form (section 6.1.2). This means that \mathbf{F} does not depend on ρ and that the third column of $\mathbf{J}_{\mathbf{F}}^{\Theta}$ is zero. In practice, if the compactness constraint is verified, we can replace in equations 6.27 to 6.42 $\Upsilon = (\phi, \psi, \rho)^t$ by $\hat{\Upsilon} = (\phi, \psi)^t$ and assume that $\mathbf{J}_{\mathbf{F}}^{\Upsilon}$ is a 2×2 matrix by discarding the null column. The verification of the compactness constraint is a necessary conditions for $\mathbf{J}_{\mathbf{F}}^{\Upsilon}$ being a square matrix with inverse. Considering the statement of proposition 6.2, the verification of the compactness constraint is a necessary condition to exist an v function. If both v and θ exist then there is an analytical solution for the active tracking problem (proposition 6.1).

From the statements above we can conclude that it is desirable to select a transformation of coordinates Γ verifying the compactness constraint (equation 6.10). If Γ also verifies the decoupling constraint (equation 6.11) then \mathbf{F} is written in a decoupled way which simplifies the calculations to obtain the tracking control laws (equations 6.34 and 6.37). Another useful guideline is to select transformation Γ such that target 3D position is referenced in a coordinate frame of the same type as the one where camera position Θ is defined. This can also lead to several simplifications in the calculation as well as a deeper understanding of the tracking task.

6.4 Active Tracking of Moving Targets

The derived mathematical framework is applied to active tracking of a moving target using three different robotic platforms. These examples help the understanding of the exposed ideas and illustrate the usefulness of a judicious selection of the world coordinate system.

6.4. Active Tracking of Moving Targets

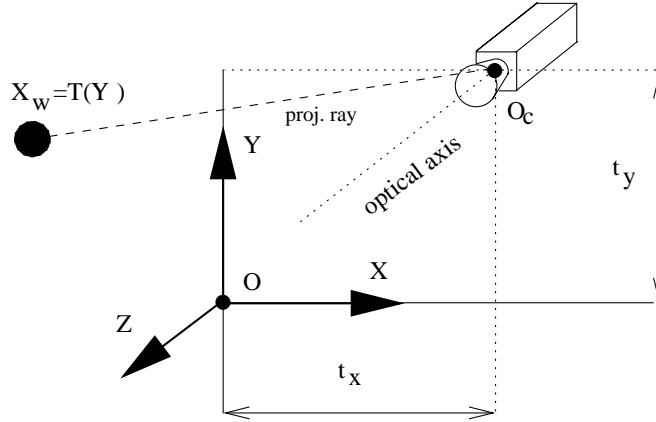


Figure 6.7: Active tracking using a perspective camera with translation motion

6.4.1 Active Tracking using a Perspective Camera with Translation Motion in the XY Plane

Fig.6.7 depicts a conventional perspective camera with translation motion in the XOY plane. A schematic of the framework to derive global mapping \mathbf{F} is shown in Fig.6.1. The imaging sensor is a perspective camera with intrinsic parameters $\mathbf{K}_c = \mathbf{I}$, therefore function f_i is the one shown in equation 6.12. The camera position in the XOY plane in world cartesian coordinates is $\Theta = (t_X, t_Y)^t$. For this particular case $\mathbf{P}(\Theta)$ is given by equation 6.43. The goal of the application is to control camera position and velocity such that target position and velocity in image are zero ($\mathbf{x}_d = \dot{\mathbf{x}}_d = 0$). Therefore the camera motion must be controlled in such a way as to keep the optical axis aligned with target mass center.

$$\mathbf{P}(\Theta) = \begin{bmatrix} 1 & 0 & 0 & -t_X \\ 0 & 1 & 0 & -t_Y \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (6.43)$$

Selecting Coordinate Transformation Γ

In a similar way to what was done in section 6.1 we intend to select a suitable coordinate transformation Γ (equation 6.5) for this specific application. Functions f_i , $\mathbf{P}(\Theta)$ and f_h have been already defined. The jacobian matrix $\mathbf{J}_F^{\mathbf{Y}}$ can be written in terms of the partial derivatives of Γ by replacing \mathbf{J}_Γ in equation 6.30 by the

6. A General Framework for Selecting the World Coordinate System

result of equation 6.6.

If we intend to select a coordinate transformation Γ verifying the compactness constraint then the third column of matrix $\mathbf{J}_\Gamma^{\mathbf{F}}$ must be zero (equation 6.10). This yields the partial differential equations shown in 6.44.

$$\begin{cases} Z(\phi_Y\psi_Z - \phi_Z\psi_Y) + (X - t_X)(\phi_Y\psi_X - \phi_X\psi_Y) = 0 \\ Z(\phi_Z\psi_X - \phi_X\psi_Z) + (Y - t_Y)(\phi_Y\psi_X - \phi_X\psi_Y) = 0 \end{cases} \quad (6.44)$$

The system of equations 6.44 can be solved with respect to the partial derivatives of Γ . \mathbf{J}_Γ is a 3×3 matrix which yields 9 unknowns for 2 equations. The problem is under determined and multiple solutions can be found. Some of the solutions obtained can be discarded by considering the additional constraint $\det(\mathbf{J}_\Gamma) \neq 0$ (transformation Γ must be bijective). Following this procedure the result of equation 6.45 is derived. If the coordinate transformation Γ verifies the compactness constraint then the structure of the corresponding jacobian matrix \mathbf{J}_Γ must be the one shown in 6.45. Notice that the provided jacobian matrix depends both on $\mathbf{X}_w = (X, Y, Z)^t$ and $\Theta = (t_x, t_y)^t$. Any function Γ verifying equation 6.44 must depend, not only on \mathbf{X}_w , but also on Θ . Γ is no longer a transformation of inertial world cartesian coordinates and we may conclude that it is not possible to find a transformation of coordinates verifying the compactness constraint.

$$\mathbf{J}_\Gamma = \begin{bmatrix} \phi_X & \phi_Y & -\frac{(X-t_X)\phi_X+(Y-t_Y)\phi_Y}{Z} \\ \psi_X & \psi_Y & -\frac{(X-t_X)\psi_X+(Y-t_Y)\psi_Y}{Z} \\ \rho_X & \rho_Y & \rho_Z \end{bmatrix} \quad (6.45)$$

The conclusions drawn from the discussion above were expected. If the camera has translation motion there is no way to suppress the dependence on the third coordinate and it is not possible to perform the required motion control using only visual information. Some authors overcome the problem by assuming additional constraints such the target moving in a plane in the world [25, 46]. A systematic approach to determine if it exists any coordinate transformation verifying the constraints specified in 6.1.2 has been presented. The proposed procedure establishes necessary conditions. The fulfillment of the conditions does not guarantee the existence of a desired coordinate transformation Γ . This systematic approach will

6.4. Active Tracking of Moving Targets

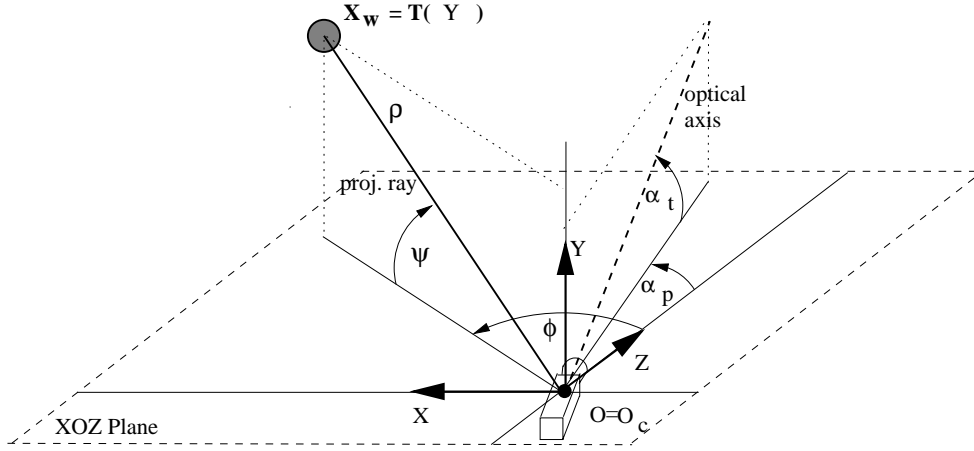


Figure 6.8: Active tracking using a perspective camera mounted on a pan and tilt unit

be repeated in the next two cases.

6.4.2 Active Tracking using a Perspective Camera with Pan and Tilt Rotation Motion

A perspective camera is mounted on a pan and tilt unit such that both rotation axes go through the optical center O_c . The camera first rotates in pan and then in tilt (Fick Model) as depicted in Fig. 6.8. The camera position vector is $\Theta = (\alpha_p, \alpha_t)^t$ with α_p the pan angle and α_t the tilt angle.

Consider the global mapping scheme depicted in Fig. 6.1. For this particular case function f_i is provided in equation 6.12 (perspective camera with $\mathbf{K}_c = \mathbf{I}$), Equation 6.46 shows $\mathbf{P}(\Theta)$ with $e^{-\alpha_p \hat{y}_c}$ and $e^{-\alpha_t \hat{x}_c}$ the pan and tilt rotation matrices [51]. Unit vectors \hat{x}_c and \hat{y}_c are associated with the X_c and Y_c axes of the cartesian coordinate frame attached to the camera (not depicted in Fig. 6.8). The goal of the tracking application is to control camera rotation such that both target position and velocity in image is zero ($\mathbf{x}_d = \dot{\mathbf{x}}_d = 0$). To achieve this the camera optical axis must be aligned with target in 3D space.

$$\mathbf{P}(\Theta) = e^{-\alpha_t \hat{x}_c} e^{-\alpha_p \hat{y}_c} [\mathbf{I} | \mathbf{0}] \quad (6.46)$$

6. A General Framework for Selecting the World Coordinate System

$$\begin{array}{l}
 \Gamma(\mathbf{X}_w) = \begin{bmatrix} \arctan(\frac{X}{Z}) \\ \arctan(-\frac{Y}{\sqrt{X^2+Z^2}}) \\ \sqrt{X^2+Y^2+Z^2} \end{bmatrix} \\
 \mathbf{T}(\Upsilon) = \begin{bmatrix} \rho \sin(\phi) \cos(\psi) \\ -\rho \sin(\psi) \\ \rho \cos(\phi) \cos(\psi) \end{bmatrix} \\
 \mathbf{F}(\Upsilon, \Theta) = \begin{bmatrix} \frac{C(\psi)S(\phi-\alpha_p)}{C(\psi)C(\alpha_t)C(\phi-\alpha_p)+S(\psi)S(\alpha_t)} \\ \frac{S(\psi)C(\alpha_t)-C(\psi)S(\alpha_t)C(\phi-\alpha_p)}{C(\psi)C(\alpha_t)C(\phi-\alpha_p)+S(\psi)S(\alpha_t)} \\ \frac{C(\psi)(S(\psi)S(\alpha_t)C(\phi-\alpha_p)+C(\psi)C(\alpha_t))}{C(\psi)S(\psi)S(\phi-\alpha_p)} \end{bmatrix} \\
 \mathbf{J}_F^\Upsilon(\Upsilon, \Theta) = \begin{bmatrix} -\frac{S(\alpha_t)S(\phi-\alpha_p)}{\Xi_1} & 0 \\ -\frac{C(\phi-\alpha_p)}{\Xi_1} & 0 \end{bmatrix} \\
 \mathbf{J}_F^\Theta(\Upsilon, \Theta) = \begin{bmatrix} \frac{C(\psi)\Xi_2}{\Xi_1} \\ \frac{\Xi_3}{\Xi_1} \end{bmatrix} \\
 \Xi_1 = C(\psi)^2C(\alpha_t)^2S(\psi)^2(S(\phi)^2+S(\alpha_p)^2+2C(\phi)C(\alpha_p)C(\phi-\alpha_p)) \\
 \quad -C(\alpha_t)^2+2S(\alpha_t)C(\alpha_t)S(\psi)C(\psi)C(\phi-\alpha_p) \\
 \Xi_2 = C(\psi)S(\alpha_t)(S(\alpha_p)C(\alpha_p)+2C(\alpha_p)C(\phi)S(\phi-\alpha_p)) \\
 \quad -S(\phi)C(\phi)-S(\psi)C(\alpha_t)S(\phi-\alpha_p) \\
 \Xi_3 = 1-C(\psi)^2(C(\phi)^2+C(\alpha_p)^2-2C(\alpha_p)C(\phi)C(\phi-\alpha_p))
 \end{array}$$

Table 6.4: Using a spherical coordinate system for active tracking with a pan and tilt perspective camera

Selecting Coordinate Transformation Γ

Functions \mathbf{f}_i , $\mathbf{P}(\Theta)$ and \mathbf{f}_h , as well as the corresponding jacobian matrices, are defined. Assume \mathbf{J}_Γ given by equation 6.6. According to equation 6.30, matrix \mathbf{J}_F^Υ can be written in terms of the partial derivatives of coordinate transformation Γ . The procedure described in section 6.4.1 is repeated. If Γ verifies the compactness constraint then its jacobian matrix must have the structure shown in equation 6.47. For the pan and tilt tracking situation there is a solution \mathbf{J}_Γ that only depends on \mathbf{X}_w . Thus the existence of a cartesian coordinate transformation Γ verifying the compactness constraint is not excluded. Nevertheless, repeating the procedure to achieve a decoupled \mathbf{F} function, allows us to conclude that it is not possible to find a transformation Γ which verifies the decoupling constraint.

6.4. Active Tracking of Moving Targets

$$\mathbf{J}_\Gamma = \begin{bmatrix} \phi_X & -\frac{X\phi_X+Z\phi_Z}{Y} & \phi_Z \\ -\frac{Y\psi_Y+Z\psi_Z}{X} & \psi_Y & \psi_Z \\ \rho_X & \rho_Y & \rho_Z \end{bmatrix} \quad (6.47)$$

One of the suggested guidelines for the selection of transformation Γ is to reference the target 3D position in a coordinate frame of the same type as the one where the camera position Θ is defined.

Consider the transformation from cartesian coordinates into spherical coordinates proposed in equation 6.48. The angles ϕ and ψ , used to reference the target 3D position, are similar to the camera pan and tilt angles (see Fig. 6.8). Moreover the jacobian matrix of the coordinate transformation 6.48 follows the structure presented in equation 6.47 which assures that the compactness constraint is verified.

$$\begin{cases} \phi = \arctan\left(\frac{X}{Z}\right) \\ \psi = \arctan\left(-\frac{Y}{\sqrt{X^2+Z^2}}\right) \\ \rho = \sqrt{X^2 + Y^2 + Z^2} \end{cases} \quad (6.48)$$

Table 6.4 summarizes the results obtained using the coordinate transformation of equation 6.48. Both Γ and its inverse \mathbf{T} are presented. Equation 6.49 defines an angular error vector Δ . The goal of the tracking application is to align the camera optical axis with the target mass center. This is verified whenever Δ is zero. Notice that by using coordinates of the same type to reference both the target and the camera position, the dependence of the global mapping \mathbf{F} on the angular tracking errors δ_p and δ_t is explicit. The third column of $\mathbf{J}_\mathbf{F}$ is zero which means that Γ verifies the compactness constraint. Due to lack of space the polynomials Ξ_1 , Ξ_2 and Ξ_3 are presented at the bottom of the table.

$$\Delta = (\delta_p, \delta_t)^t = (\phi - \alpha_p, \psi - \alpha_t)^t \quad (6.49)$$

Active Tracking Control Law

Assume that Θ_c and $\dot{\Theta}_c$ are the position and velocity commands that must be sent to platform actuators to accomplish a specified task (the actuators transfer function is unitary). The goal is to track a moving object such that its projection

6. A General Framework for Selecting the World Coordinate System

$$\begin{array}{l}
 \Theta_c = \Theta + \\
 \dot{\Theta}_c = \dot{\Theta} +
 \end{array}
 \left[\begin{array}{c}
 \arctan\left(\frac{x_i}{y_i S(\alpha_t) + C(\alpha_t)}\right) \\
 \arctan\left(\frac{\frac{y_i - \tan(\alpha_t)}{1 + y_i \tan(\alpha_t)} + \tan(\alpha_t) \sqrt{1 + \frac{x_i^2}{(y_i S(\alpha_t) + C(\alpha_t))^2}}}{\frac{y_i - \tan(\alpha_t)}{1 + y_i \tan(\alpha_t)} \tan(\alpha_t) - \sqrt{1 + \frac{x_i^2}{(y_i S(\alpha_t) + C(\alpha_t))^2}}}\right) \\
 \frac{1 + y_i^2}{(1 + x_i^2 + y_i^2)(y_i S(\alpha_t) + C(\alpha_t))} \quad - \frac{x_i y_i}{(1 + x_i^2 + y_i^2)(y_i S(\alpha_t) + C(\alpha_t))} \\
 \frac{x_i (y_i C(\alpha_t) - S(\alpha_t))}{(1 + x_i^2 + y_i^2)(y_i S(\alpha_t) + C(\alpha_t))} \quad - \frac{(y_i S(\alpha_t) + (1 + x_i^2) C(\alpha_t))}{(1 + x_i^2 + y_i^2)(y_i S(\alpha_t) + C(\alpha_t))}
 \end{array} \right] \dot{\mathbf{x}}_i$$

Table 6.5: Position and velocity command for active tracking with a pan and tilt perspective camera

is kept in the image center. The position command Θ_c can be obtained by making $\mathbf{x}_d = \mathbf{0}$ in equation 6.34 ($\Theta_c = \theta(\mathbf{0}, v(\mathbf{x}_i, \Theta))$). Notice that the specified tracking task is accomplished by controlling the camera motion such that the optical axis becomes aligned with target mass center. Therefore the position command must be $\Theta_c = \hat{\mathbf{Y}}$ with $\hat{\mathbf{Y}} = (\phi, \psi)^t$. Considering that $\hat{\mathbf{Y}} = \Theta + \Delta$ (equation 6.49) it results $\Theta_c = \Theta + \Delta$. The explicit computation of v and θ functions can be avoided by determining angular error Δ as a function of target position in image \mathbf{x}_i and camera pose Θ

In table 6.4 the global mapping \mathbf{F} is written in a compact form such that \mathbf{x}_i depends on $\hat{\mathbf{Y}}$ and Θ . The 3D position parameters ϕ and ψ can be recovered from the target position in the image (x_i, y_i) and the camera pose (α_p, α_t) (equation 6.50). The derivation of the angular error vector Δ from equation 6.48 as a function of target position in the image and the camera pose is straightforward. The position command Θ_c is obtained by making $\Theta_c = \Theta + \Delta$ (table 6.5).

$$\begin{cases}
 \phi = \arctan\left(\frac{(y_i \sin(\alpha_t) + \cos(\alpha_t)) \sin(\alpha_p) + x_i \cos(\alpha_p)}{(y_i \sin(\alpha_t) + \cos(\alpha_t)) \cos(\alpha_p) - x_i \sin(\alpha_p)}\right) \\
 \psi = -\arctan\left(\frac{y_i \cos(\alpha_t) - \sin(\alpha_t)}{\sqrt{x_i^2 + (y_i \sin(\alpha_t) + \cos(\alpha_t))^2}}\right)
 \end{cases} \quad (6.50)$$

Consider the jacobian matrices of table 6.4. Replacing angles ϕ and ψ by the result of equation 6.50, both $\mathbf{J}_F^{\mathbf{Y}}$ and \mathbf{J}_F^{Θ} can be written in terms of target position in the image and the camera pose. Equation 6.51 shows $\hat{\mathbf{J}}_F^{\mathbf{Y}}$ such that $\mathbf{J}_F^{\mathbf{Y}} = [\hat{\mathbf{J}}_F^{\mathbf{Y}} | \mathbf{0}]$, and equation 6.52 gives \mathbf{J}_F^{Θ} . Making $\dot{\mathbf{x}}_d = \mathbf{0}$ in equation 6.42 the velocity command $\dot{\Theta}$ is obtained (table 6.5).

6.4. Active Tracking of Moving Targets

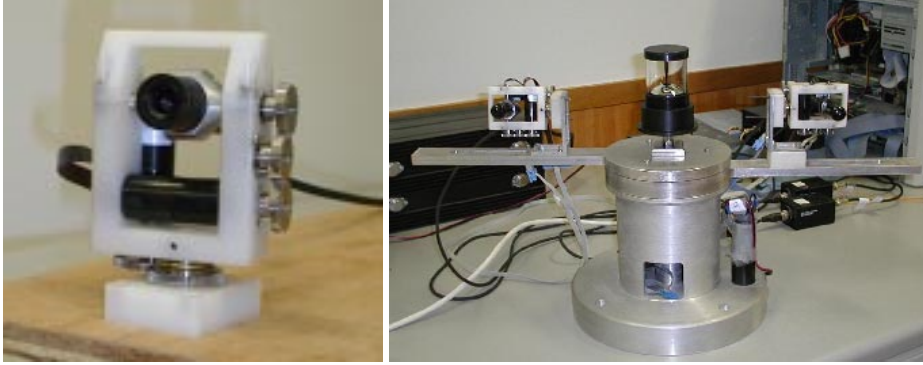


Figure 6.9: The Modular Vision System (MVS) at ISR

$$\hat{\mathbf{J}}_{\mathbf{F}}^{\Upsilon} = \begin{bmatrix} y_i S(\alpha_t) + (1 + x_i^2)C(\alpha_t) & -\frac{x_i S(\alpha_t)(1 + x_i^2 + y_i^2)}{\sqrt{x_i^2 + (y_i S(\alpha_t) + C(\alpha_t))^2}} \\ (y_i C(\alpha_t) - S(\alpha_t))x_i & -\frac{(y_i S(\alpha_t) + C(\alpha_t))(1 + x_i^2 + y_i^2)}{\sqrt{x_i^2 + (y_i S(\alpha_t) + C(\alpha_t))^2}} \end{bmatrix} \quad (6.51)$$

$$\mathbf{J}_{\mathbf{F}}^{\Theta} = \begin{bmatrix} -(y_i S(\alpha_t) + (1 + x_i^2)C(\alpha_t)) & x_i y_i \\ -(y_i C(\alpha_t) - S(\alpha_t))x_i & 1 + y_i^2 \end{bmatrix} \quad (6.52)$$

If $y_i = -\cot(\alpha_t)$ then $y_i \sin(\alpha_t) + \cos(\alpha_t) = 0$ and matrix $\mathbf{J}_{\mathbf{F}}^{\Theta}$ becomes non-invertible and a singularity occurs in the derived expressions of Θ_c and $\dot{\Theta}_c$. This happens whenever target image lays in an horizontal line that contains the intersection point of pan rotation axis with image plane. For this case $\delta_p = \pm\pi/2$ and $\tan(\delta_p) = \pm\infty$.

6.4.3 Tracking Applications Using the MDOF and the MVS Robotic platforms

In our laboratory two robotic platforms have been developed for active tracking: the MDOF robot head [16] and the MVS modular platform. The MDOF head has been mainly used in monocular and binocular tracking of a single target [17, 15, 14]. The MVS head has been recently built with the purpose of working on simultaneous tracking of multiple targets [13]. Fig. 6.9(Left) depicts one of the MVS robotic eyes. The camera has two degrees of freedom: pan and tilt. The

6. A General Framework for Selecting the World Coordinate System

rotation axes go through the optical center and the camera undergoes pure rotation motion. This section describes the application of the equations derived in section 6.4.2 for monocular tracking. Two different behaviors must be considered: the saccadic motion and the smooth pursuit [16, 17].

The active vision system is in standby until the target appears in the field of view. Initially the object is projected somewhere in the image periphery. The goal of the saccadic motion is to position the target image in the foveal region. The performance of the saccadic control is highly dependent on the accuracy of the angular error estimation. The importance of using the exact position command of Tab. 6.5 to accomplish the task is shown. After a successful saccadic motion the object is projected nearby the image center. The smooth pursuit behavior adjusts the system position such that the target is kept in the image center. We show that for the smooth pursuit the global mapping \mathbf{F} can be approximated by function $\bar{\mathbf{F}}$ which yields simpler mathematical expressions and decoupled control of pan and tilt DOF. Function $\bar{\mathbf{F}}$ is established assuming that the target is projected near the image center during tracking. The approximation errors are studied.

Saccadic Motion

As mentioned, the angular error $\Delta = (\delta_p, \delta_t)^t$ (equation 6.49) can be written as a function of the target position in the image and of the camera pose (table 6.5). Solving the system of equations with respect to x_i and y_i yields the result of equation 6.53.

$$\begin{cases} x_i = (y_i S(\alpha_t) + C(\alpha_t)) \tan(\delta_p) \\ y_i = -\frac{(S(\alpha_t)^2 C(\delta_p) + C(\alpha_t)^2) \tan(\delta_t) + S(\alpha_t) C(\alpha_t) (1 - C(\delta_p))}{(S(\alpha_t)^2 + C(\alpha_t)^2 C(\delta_p)) + S(\alpha_t) C(\alpha_t) (1 - C(\delta_p)) \tan(\delta_t)} \end{cases} \quad (6.53)$$

The saccadic motion consists in rotating the camera such that target image jumps from the periphery to the center of the retina. Tab. 6.5 provides the exact position command Θ_c to accomplish the task. If command Θ_c is sent to pan and tilt motors, then the angular error vector Δ becomes null. The saccadic motion is perfect because target projection moves from its initial position on the image periphery to the center (replace δ_p, δ_t by zero in equation 6.53).

The results presented on Tab. 6.5 are novel. In [24] the equations to track an object moving in a plane with known depth are derived. The control equations

6.4. Active Tracking of Moving Targets

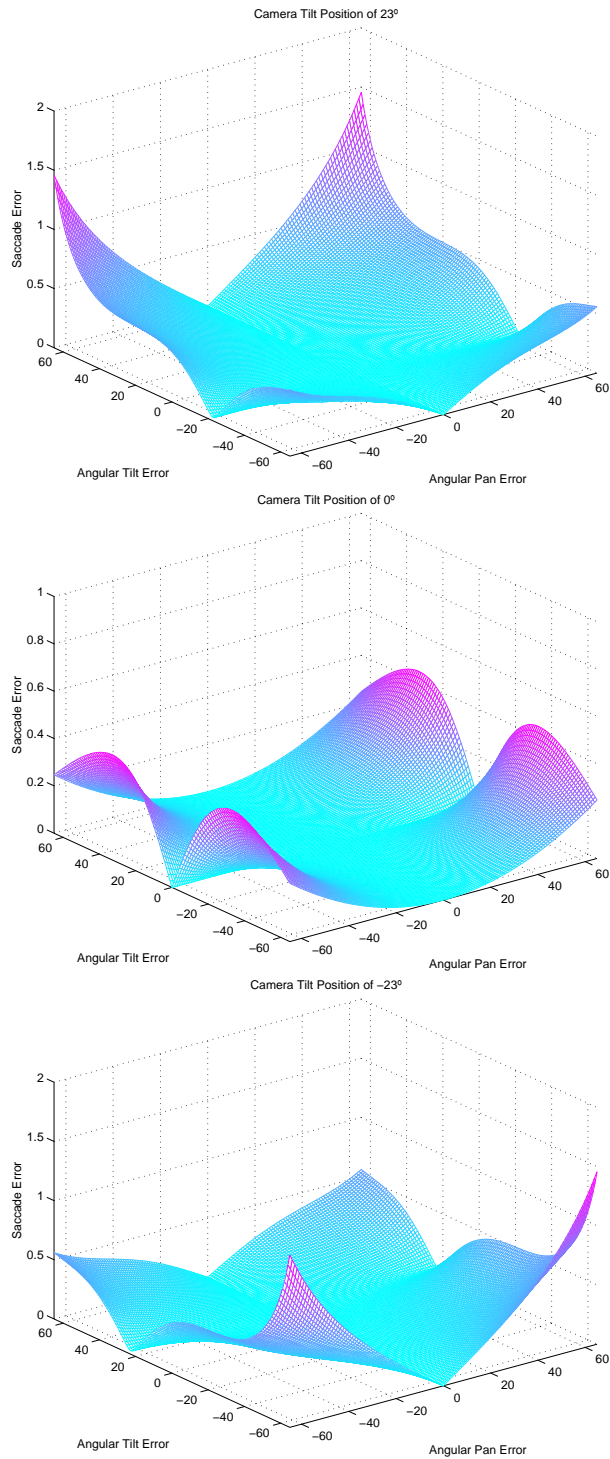


Figure 6.10: Distance to the image center after the saccadic motion.

6. A General Framework for Selecting the World Coordinate System

proposed on Tab. 6.5 generalize this result since they assume unconstrained target motion. In [60, 17] the estimation of the angular error Δ is approximated by $\bar{\Delta} = (\arctan(x_i), -\arctan(y_i))^t$. This approximation is rather intuitive and simplifies the derivation of the control law. However the performance of the saccadic control tends to be poor as shown in Fig. 6.10. Consider that the position command is $\bar{\Theta}_c = \Theta + \bar{\Delta}$. After the saccadic motion, target image moves towards the foveal region since angles δ_p and δ_t decrease. However the error angles do not become necessarily zero, nor the target is projected in the image center. Fig. 6.10 shows the distance to the image center after the saccadic motion for three different camera tilt positions. The X and Y axes correspond to the initial pan and tilt errors (δ_p, δ_t) which take values in the range $[-60^\circ, 60^\circ]$. The Z axis is the normalized distance to the image center after the saccadic motion. To convert into image pixels the value must be multiplied by $k.f$ with k the number of pixels per metric unit and f the camera focal length (in our case $k.f = 225$ for both the MDOF and MVS robot head).

Observing Fig. 6.10 comes that the performance of the saccadic motion decreases when the angular errors increase. Since the target starts by appearing in the image periphery, both δ_p and δ_t take in general high values. One concludes that the performance of the saccadic control using command $\bar{\Theta}_c$ tends to be poor. The approximation $\bar{\Delta} = (\arctan(x_i), -\arctan(y_i))^t$ is only valid for small angular errors and camera tilt angles. The command equation of Tab. 6.5 must be used to control the saccadic motion.

Function \bar{F} for Smooth Pursuit

In general the active tracking process is initialized by the saccadic motion. Camera pose changes abruptly such that target projection jumps from the image periphery to the foveal region. After the saccadic motion the camera orientation is smoothly adjusted to keep the target image in the center of the retina. This stage is called the smooth pursuit control.

The global mapping function F is shown in table 6.4. In the smooth pursuit it is reasonable to assume that most of the time the target projection is near the image center. This assumption is used to derive an approximate mapping \bar{F} . Equation 6.54 results from making $y_i = 0$ in the first equation of 6.53 and $\delta_p = 0$ in

6.4. Active Tracking of Moving Targets

the second one. Function $\bar{\mathbf{F}}$ is an approximation of the global mapping \mathbf{F} . The approximation has simpler mathematical expressions and presents the advantage of decoupling the pan and tilt control.

$$\bar{\mathbf{F}}(\Upsilon, \Theta) = \begin{bmatrix} \cos(\alpha_t) \tan(\delta_p) \\ -\tan(\delta_t) \end{bmatrix} \quad (6.54)$$

Consider the approximation error vector $\mathbf{E} = \mathbf{x}_i - \bar{\mathbf{x}}_i$ (equation 6.55) measured in the image plane. The exact image target position is $\mathbf{x}_i = \mathbf{F}(\Upsilon, \Theta)$, and $\bar{\mathbf{x}}_i = \bar{\mathbf{F}}(\Upsilon, \Theta)$ is the approximate one. Error \mathbf{E} depends on the camera tilt position α_t and on the angular tracking error Δ .

$$\mathbf{E}(\alpha_t, \Delta) = (E_x(\alpha_t, \delta_p, \delta_t), E_y(\alpha_t, \delta_p, \delta_t))^t \quad (6.55)$$

Fig. 6.11 helps to understand the approximation function $\bar{\mathbf{F}}$. Each figure represents the image plane for three different camera tilt positions. If α_t is known then target position in the image depends on the angular tracking error Δ . Both the exact (\mathbf{x}_i) and approximate ($\bar{\mathbf{x}}_i$) target position in the image are calculated for a set of values of Δ . The values computed of \mathbf{x}_i and $\bar{\mathbf{x}}_i$ are represented in the image plane. The exact positions on the solid grid are approximated by the positions of the dashed grid. Conclusions about $\mathbf{E} = (E_x, E_y)^t$ can be drawn by observing Fig. 6.11.

Assume that the angular pan error δ_p is constant along time. The target is positioned somewhere in a vertical plane in the 3D world, going through the origin O of the inertial coordinate frame. This plane is projected in a line in the image. If $\alpha_t = 0$, the line is vertical, if $\alpha_t \neq 0$ the line has a slope whose module is inversely proportional to module of camera tilt angle. Using the approximation of equation 6.54 the mentioned plane, containing the target, is projected in a vertical line in the image. As depicted in Fig. 6.11, $\bar{x}_i = x_i$ whenever $\alpha_t = 0$ or $y = 0$. Some properties of error function $E_x(\alpha_t, \delta_p, \delta_t)$ can be observed:

- $E_x(0, \delta_p, \delta_t) = 0$
- $E_x(\alpha_t, \delta_p, \delta_t) = -E_x(\alpha_t, -\delta_p, \delta_t)$
- $E_x(\alpha_t, \delta_p, \delta_t) = E_x(-\alpha_t, \delta_p, -\delta_t)$

6. A General Framework for Selecting the World Coordinate System

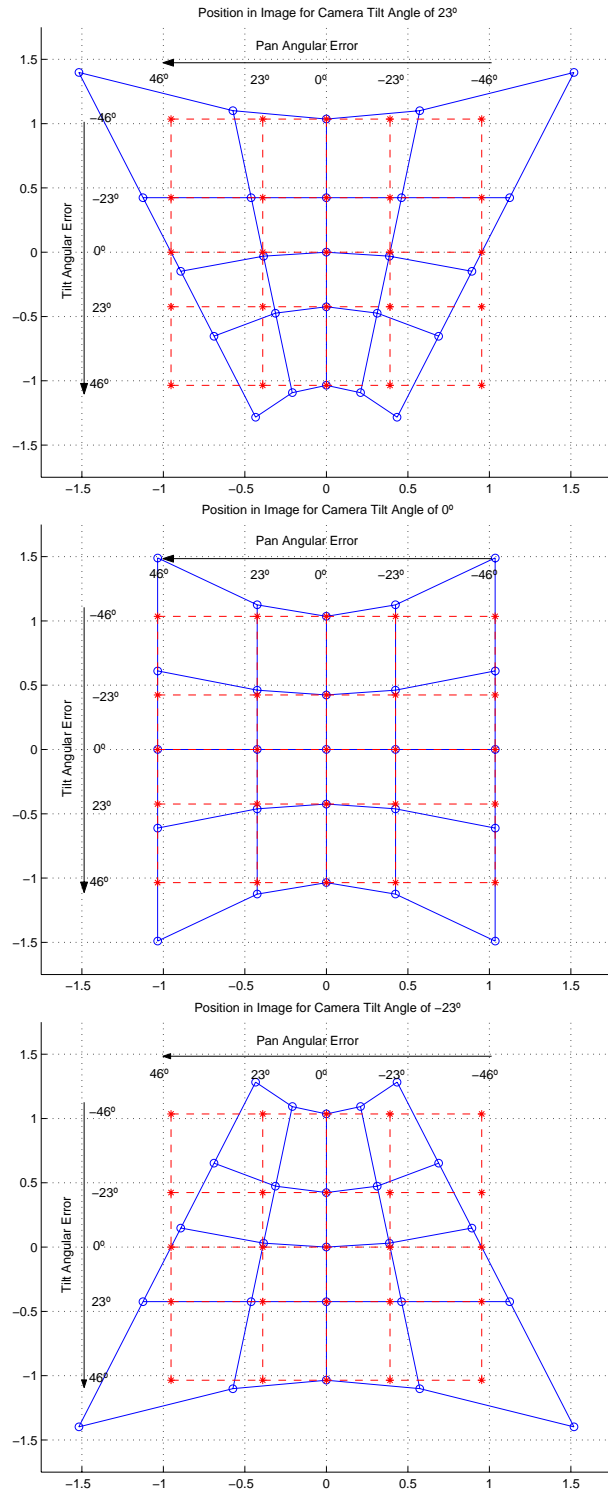


Figure 6.11: Approximating the global mapping function F by \bar{F} . Exact (o) and approximated (*) target position.

6.4. Active Tracking of Moving Targets

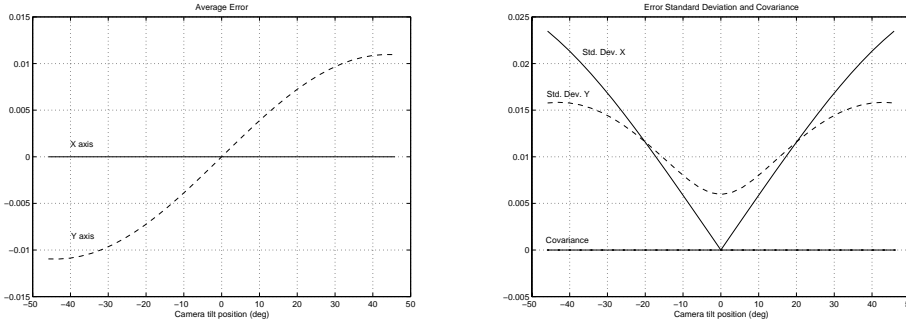


Figure 6.12: Statistical analysis of approximation error function \mathbf{E} . Left: Mean of E_x (-) and E_y (- -). Right: Standard deviation of E_x (-) and E_y (- -).

Consider that the angular tilt error δ_t constant. In 3D space the target is somewhere in a cone with vertex in the origin O of the inertial coordinate frame. The image projection of these surfaces are the hyperbolic lines shown in Fig. 6.11. Whenever $\delta_t = -\alpha_t$ the conic surface degenerates in the OXZ plane. The plane projection is an horizontal line also observable in the figure. The approximation of equation 6.54 generates horizontal lines in the image. Notice that $\bar{y}_i = y_i$ whenever $\delta_t = -\alpha_t$ or $\delta_p = 0$. The following properties of error function $E_y(\alpha_t, \delta_p, \delta_t)$ can be observed:

- $E_y(\alpha_t, 0, \delta_t) = 0$
- $E_y(\alpha_t, \delta_p, -\alpha_t) = 0$
- $E_y(\alpha_t, \delta_p, \delta_t) = E_y(\alpha_t, -\delta_p, \delta_t)$
- $E_y(\alpha_t, \delta_p, \delta_t) = -E_y(-\alpha_t, -\delta_p, -\delta_t)$

A statistical characterization of the angular tracking error $\mathbf{\Delta} = (\delta_p, \delta_t)^t$ has been done experimentally. Assume that both δ_p and δ_t have an independent gaussian probability distribution of average 0° (the target is almost always near the center of the image). The standard deviation of δ_p is 12° and the standard deviation of δ_t is 8° . The pan and tilt errors are statistically independent thus the covariance is zero. The approximation errors \mathbf{E}_x and \mathbf{E}_y (equation 6.55) depend on α_t and $\mathbf{\Delta}$. Given the camera tilt position α_t and the statistical characterization

6. A General Framework for Selecting the World Coordinate System

of the angular tracking error Δ , the statistical properties of E_x and E_y can be derived. Fig. 6.12 shows the averages μ_x and μ_y , the standard deviations σ_x and σ_y and the covariance σ_{xy} as a function of camera tilt angle α_t . μ_x and σ_{xy} are zero because E_x is an odd function of δ_p . μ_y is an odd function of α_t because $E_y(\alpha_t, \delta_p, \delta_t) = -E_y(-\alpha_t, -\delta_p, -\delta_t)$. The figure presents normalized values for the statistical parameters. To convert into image pixels the value must be multiplied $k.f$ with k the number of pixels per metric unit and f the camera focal length (in our case $k.f = 225$ for both the MDOF and MVS robot head).

The approximation error E tends to increase with the tracking angular errors Δ . E also increases with the absolute value of the camera tilt angle. If $|\alpha_t|$ is high then significative approximation errors arise even when the target is projected near the image center. If we intend to replace the global mapping F by function \bar{F} then both Δ and α_t must be small, which means that the target image must be near the center and the operating range of tilt DOF can not be large.

Velocity Relationships

From equation 6.29 results that the target velocity in the image is the sum $\dot{\mathbf{x}}_i = \dot{\mathbf{x}}_{\text{ind}} + \dot{\mathbf{x}}_{\text{ego}}$ with $\dot{\mathbf{x}}_{\text{ind}} = \mathbf{J}_F^Y \dot{\mathbf{Y}}$ the velocity induced by target motion, and $\dot{\mathbf{x}}_{\text{ego}} = \mathbf{J}_F^\Theta \dot{\Theta}$ the velocity induced by camera motion (egomotion). Fig.6.13 depicts the image velocity field for different circumstances. The columns correspond to different camera tilt angles α_t (from left to right $\alpha_t = -23^\circ, 0^\circ, 23^\circ$). The first two rows depict the image velocity fields $\dot{\mathbf{x}}_{\text{ego}}$ when the camera moves in pan and in tilt and the target is stopped ($\dot{\mathbf{Y}} = (0, 0)^t$). In the first row the camera velocity is $\dot{\Theta} = (1, 0)^t$, while in the second is $\dot{\Theta} = (0, 1)^t$. The third row shows the image velocity field $\dot{\mathbf{x}}_i$ when both the camera and the target move ($\dot{\Theta} = (1, 1)^t$ and $\dot{\mathbf{Y}} = (1, 1)^t$). It is assumed that the camera has a field of view (FOV) of 86° . The small dashed rectangle corresponds to a FOV of $24^\circ \times 18^\circ$.

The goal of our tracking application is to keep a zero target image velocity $\dot{\mathbf{x}}_i = 0$. Considering that $\dot{\mathbf{x}}_i = \dot{\mathbf{x}}_{\text{ind}} + \dot{\mathbf{x}}_{\text{ego}}$, the camera velocity must be $\dot{\Theta}_c$ such that the egomotion compensates for the image velocity component induced by target motion in the scene ($\dot{\mathbf{x}}_{\text{ego}} = -\dot{\mathbf{x}}_{\text{ind}}$). The velocity command $\dot{\Theta}_c$ is written in equation 6.56 as a function of the target velocity in the world (the compactness constraint is verified, therefore $\dot{\rho}$ does not play any role). Notice that if $\dot{\psi} = 0$

6.4. Active Tracking of Moving Targets

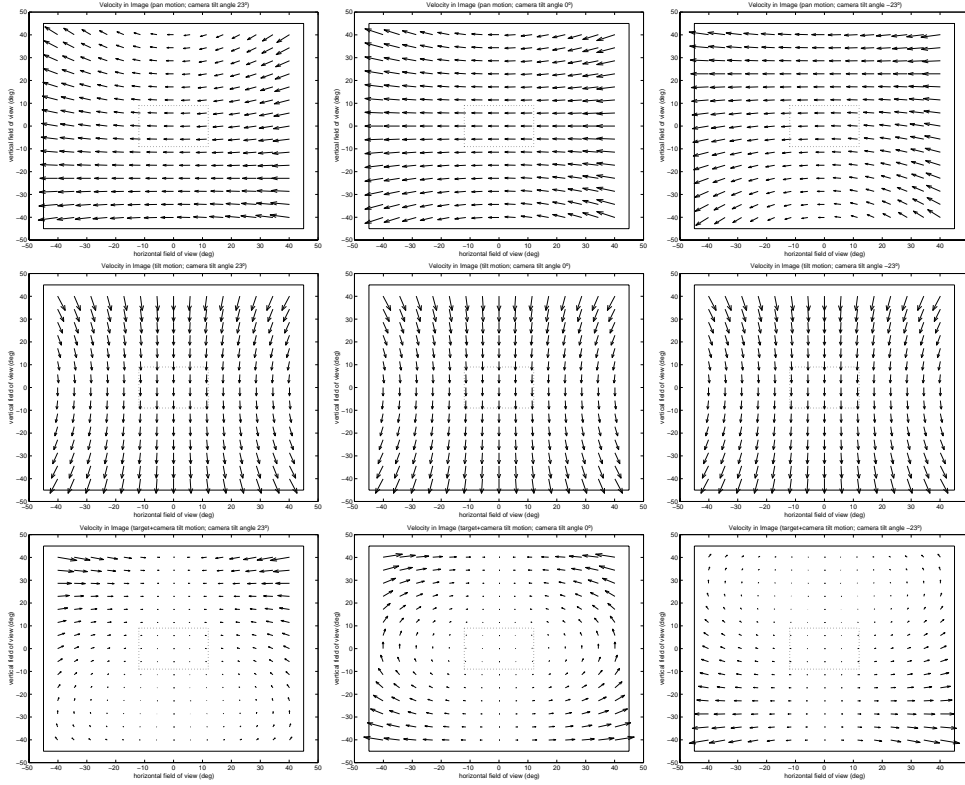


Figure 6.13: Image velocity field for different situations

then perfect tracking is achieved for $\dot{\Theta}_c = (\dot{\phi}, 0)^t$. The image velocity induced by the target motion is compensated for by the camera pan rotation. In general the decoupling between pan and tilt control does not hold as can be seen in the last row of Fig 6.13.

$$\begin{aligned} \dot{\Theta}_c &= -(\mathbf{J}_F^\Theta)^{-1} \hat{\mathbf{J}}_F^\Upsilon \begin{bmatrix} \dot{\phi} \\ \dot{\psi} \end{bmatrix} \\ &= \begin{bmatrix} 1 & \frac{x_i(y_i C(\alpha_t) - S(\alpha_t))}{(y_i S(\alpha_t) + C(\alpha_t)) \sqrt{x_i^2 + (y_i S(\alpha_t) + C(\alpha_t))^2}} \\ 0 & \frac{\sqrt{x_i^2 + (y_i S(\alpha_t) + C(\alpha_t))^2}}{y_i S(\alpha_t) + C(\alpha_t)} \end{bmatrix} \begin{bmatrix} \dot{\phi} \\ \dot{\psi} \end{bmatrix} \end{aligned} \quad (6.56)$$

The jacobian matrices $\hat{\mathbf{J}}_F^\Upsilon$ and \mathbf{J}_F^Θ can be approximated by $\bar{\mathbf{J}}_F^\Upsilon$ and $\bar{\mathbf{J}}_F^\Theta$. The result of equation 6.57 is derived assuming in equations 6.51 and 6.52 that the

6. A General Framework for Selecting the World Coordinate System

target is projected in the center of the image ($\mathbf{x}_i = 0$). The approximated jacobian matrices yield simpler mathematical expressions and decoupled pan and tilt control. The egomotion term is approximated by $\dot{\mathbf{x}}_{\text{ego}} = (\cos(\alpha_t)\dot{\alpha}_p, \dot{\alpha}_t)^t$ which is independent of the image position. As can be observed in Fig. 6.13 the egomotion is only constant around the image center. The approximated velocity command $\dot{\Theta}_c$ is derived making $\mathbf{x}_i = 0$ in equation 6.56. As a result $\dot{\Theta}_c = (\dot{\phi}, \dot{\psi})^t$ with the target world velocity approximated by $(\dot{\phi}, \dot{\psi})^t = (\dot{x}_i / \cos(\alpha_t), \dot{y}_i)^t$. As can be seen in Fig. 6.13 the approximation yielding decoupled pan and tilt control is only valid for the central area of the image.

$$\bar{\mathbf{J}}_F^{\mathbf{Y}} = -\bar{\mathbf{J}}_F^{\Theta} = \begin{bmatrix} \cos(\alpha_t) & 0 \\ 0 & -1 \end{bmatrix} \quad (6.57)$$

6.4.4 Active Tracking Using an Omnidirectional Camera with a Rotational Degree of Freedom Around the Z axis

Fig. 6.14 shows a picture of the MVS vision system. The robotic eyes of Fig. 6.9(Left) are mounted on the tips of a rotative platform. We intend to use the omnidirectional camera to control the platform rotation. The catadioptric system is mounted in the center of the robotic platform (Fig. 6.9(Right)). The platform rotation axis goes through the catadioptric effective viewpoint \mathbf{O} and the re-projection center \mathbf{O}_c . The goal of the tracking applications is to control the rotation angle α , such that the target image lays in the Y axis of the catadioptric image plane. Fig. 6.14 represents the setup.

Function \mathbf{f}_i on the global mapping scheme of Fig. 6.1 is provided in equation 6.19 (our catadioptric system is parabolic thus $\xi = 1$). $\mathbf{P}(\Theta)$ is given in equation 6.58 where $e^{-\alpha\hat{\mathbf{z}}}$ is the matrix rotation around the Z axis (Fig. 6.14) [51]. In this case vector Θ has 1×1 dimension ($\Theta = \alpha$). In the sequel vectors Θ and Θ_c will be replaced by α and α_c to reference the camera pose and command. The goal of the tracking application is to control the camera rotation such that the X target image position and velocity are zero. The desired image target position is $\mathbf{x}_d = (0, y_d)$ and the velocity is $\dot{\mathbf{x}}_d = (0, \dot{y}_d)$ with y_d and \dot{y}_d arbitrary values.

$$\mathbf{P}(\Theta) = e^{-\alpha\hat{\mathbf{z}}}[\mathbf{I}|\mathbf{0}] \quad (6.58)$$

6.4. Active Tracking of Moving Targets

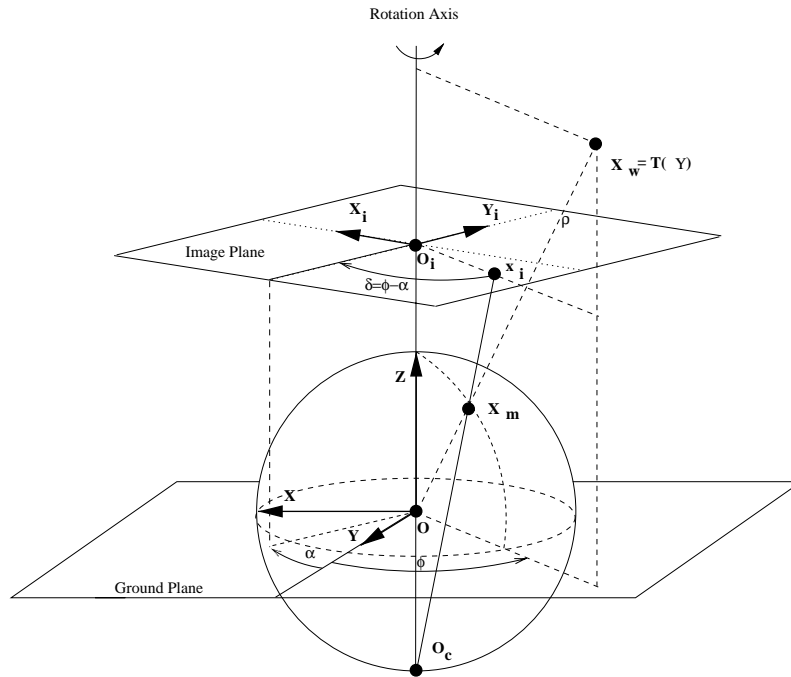


Figure 6.14: Active tracking using an omnidirectional camera

Selecting Coordinate Transformation Γ

The procedure explained in sections 6.4.1 is repeated to select a suitable world coordinate frame to reference the target position.

For this particular case it is possible to choose a coordinate transformation Γ verifying the compactness constraint. Nevertheless it is not possible to write global mapping F in a decoupled form.

The proposed change of coordinates is given in equation 6.59. It is similar to the coordinate transformation used in static catadioptric imaging (equation 6.23). The only difference is in the ϕ coordinate which is defined according to the camera position parameter α . The projective ray constraining the target is defined by the intersection of a vertical plane, referenced by ϕ , with an horizontal conic surface, indexed by ψ . The proposed coordinate system can not be used to reference points lying on the XOZ plane (Fig. 6.14). We assume that the target is always in front of the MVS system ($Y > 0$). In this case the transformation of equation 6.59 is bijective.

6. A General Framework for Selecting the World Coordinate System

$$\begin{array}{l}
 \mathbf{\Gamma}(\mathbf{X}_w) = \begin{bmatrix} \arctan(-\frac{X}{Y}) \\ \arctan(-\frac{Y}{Z+\xi\sqrt{X^2+Y^2+Z^2}}) \\ \sqrt{X^2+Y^2+Z^2} \end{bmatrix} \\
 \mathbf{T}(\Upsilon) = \begin{bmatrix} \frac{\tan(\phi)\tan(\psi)(l+\sqrt{1+(1-\xi^2)\tan(\psi)^2(1+\tan(\phi)^2)})}{1+\tan(\psi)^2(1+\tan(\phi)^2)} \\ -\frac{(\tan(\psi)(l+\sqrt{1+(1-\xi^2)\tan(\psi)^2(1+\tan(\phi)^2)}))}{1+\tan(\psi)^2(1+\tan(\phi)^2)} \\ \frac{l(1-\tan(\psi)^2(1+\tan(\phi)^2))+\sqrt{1+(1-\xi^2)\tan(\psi)^2(1+\tan(\phi)^2)}}{1+\tan(\psi)^2(1+\tan(\phi)^2)} \end{bmatrix} \\
 \mathbf{F}(\Upsilon, \alpha) = (\varphi - \xi) \cos(\alpha) \tan(\psi) \begin{bmatrix} \tan(\phi) - \tan(\alpha) \\ 1 + \tan(\phi) \tan(\alpha) \end{bmatrix} \\
 \mathbf{J}_F^\Upsilon(\Upsilon, \alpha) = (\varphi - \xi) \cos(\alpha) \begin{bmatrix} \frac{\tan(\psi)}{\cos(\phi)^2} & \frac{\tan(\phi) - \tan(\alpha)}{\cos(\psi)^2} & 0 \\ \frac{\tan(\psi)\tan(\alpha)}{\cos(\phi)^2} & \frac{1 + \tan(\phi)\tan(\alpha)}{\cos(\psi)^2} & 0 \end{bmatrix} \\
 \mathbf{J}_F^\alpha(\Upsilon, \alpha) = -(\varphi - \xi) \cos(\alpha) \tan(\psi) \begin{bmatrix} 1 + \tan(\phi) \tan(\psi) \\ -(\tan(\phi) - \tan(\alpha)) \end{bmatrix}
 \end{array}$$

Table 6.6: New coordinate system for active tracking with an omnidirectional camera

$$\begin{cases} \phi = \arctan(-\frac{X}{Y}) \\ \psi = \arctan(-\frac{Y}{Z+\xi\sqrt{X^2+Y^2+Z^2}}) \\ \rho = \sqrt{X^2+Y^2+Z^2} \end{cases} \quad (6.59)$$

Table 6.6 summarizes the results obtained. The coordinate transformation $\mathbf{\Gamma}$ and its inverse \mathbf{T} are presented. The global mapping \mathbf{F} is written using the new world coordinates $\Upsilon = (\phi, \psi, \rho)^t$. The corresponding jacobian matrices are presented as well. Notice that the third column of \mathbf{J}_F^Υ is zero, which means that transformation $\mathbf{\Gamma}$ verifies the compactness constraint.

Active Tracking Control Law

Consider the angular tracking error $\delta = \phi - \alpha$. The result of equation 6.60 is derived in a straightforward manner and provides the angular position error δ as a function of target image coordinates.

$$\delta = \arctan\left(\frac{x_i}{y_i}\right) \quad (6.60)$$

6.4. Active Tracking of Moving Targets

The global mapping function \mathbf{F} is in a compact form, thus it does not depend on ρ . Solving the equation $\mathbf{x}_i = \mathbf{F}(\Upsilon, \alpha)$ with respect to Υ , both ϕ and ψ are obtained as a function of target position in the image and catadioptric system pose. Replacing the result in the two first columns of \mathbf{J}_F^Υ the compact jacobian $\hat{\mathbf{J}}_F^\Upsilon$ is obtained as a function of image coordinates and camera pose (equation 6.61). Matrix \mathbf{J}_F^α of equation 6.62 is derived in the same way. The singularity for $\frac{x_i}{y_i} = \cot(\alpha)$ is similar to the one that appears for the pan and tilt tracking situation.

$$\hat{\mathbf{J}}_F^\Upsilon = \begin{bmatrix} \frac{x_i^2 + y_i^2}{y_i - x_i \tan(\alpha)} & \frac{((\varphi - \xi)^2 (1 + \tan(\alpha)^2) + (y_i - x_i \tan(\alpha))^2) x_i \cos(\alpha)}{(\varphi - \xi)(y_i - x_i \tan(\alpha))} \\ \frac{\tan(\alpha)(x_i^2 + y_i^2)}{y_i - x_i \tan(\alpha)} & \frac{((\varphi - \xi)^2 (1 + \tan(\alpha)^2) + (y_i - x_i \tan(\alpha))^2) y_i \cos(\alpha)}{(\varphi - \xi)(y_i - x_i \tan(\alpha))} \end{bmatrix} \quad (6.61)$$

$$\mathbf{J}_F^\alpha = (-y_i, x_i)^t \quad (6.62)$$

The X coordinate of the target position in image is zero whenever the angular error δ is zero. The position command such that x_i becomes null is $\alpha_c = \alpha + \delta$ with δ provided by equation 6.60.

Consider the target velocity vector $\dot{\Upsilon} = (\dot{\phi}, \dot{\psi})^t$. Since matrix $\hat{\mathbf{J}}_F^\Upsilon$ is non-singular, comes that $\dot{\Upsilon} = (\hat{\mathbf{J}}_F^\Upsilon)^{-1}(\dot{\mathbf{x}}_i - \mathbf{J}_F^\alpha \dot{\alpha})$ (equations 6.35 and 6.40). Assume that $\bar{\mathbf{J}}_F^\Upsilon$ and $\bar{\mathbf{J}}_F^\alpha$ are the first rows of $\hat{\mathbf{J}}_F^\Upsilon$ and \mathbf{J}_F^α . The goal is to determine the velocity command $\dot{\alpha}_c$ such that the image velocity \dot{x}_i , along the X direction, is zero. The desired velocity is $\dot{x}_d = \bar{\mathbf{J}}_F^\Upsilon \dot{\Upsilon} + \bar{\mathbf{J}}_F^\alpha \dot{\alpha}_c$ (equation 6.29). Making $\dot{x}_d = 0$ and replacing $\dot{\Upsilon}$ comes that $\dot{\alpha}_c = -(\bar{\mathbf{J}}_F^\alpha)^{-1} \bar{\mathbf{J}}_F^\Upsilon (\hat{\mathbf{J}}_F^\Upsilon)^{-1} (\dot{\mathbf{x}}_i - \mathbf{J}_F^\alpha \dot{\alpha})$. The result of equation 6.63 is achieved taking into account that $\bar{\mathbf{J}}_F^\Upsilon (\hat{\mathbf{J}}_F^\Upsilon)^{-1} = (1, 0)$.

$$\begin{cases} \alpha_c = \alpha + \arctan\left(\frac{x_i}{y_i}\right) \\ \dot{\alpha}_c = \dot{\alpha} - \frac{\dot{x}_i}{y_i} \end{cases} \quad (6.63)$$

The control law of equation 6.63 allows to control the angular position and velocity of the platform where the catadioptric sensor is mounted on in a straightforward manner.

6.5 Closure

Any image formation process can be interpreted as a transformation from \mathbf{R}^3 to \mathbf{R}^2 . The transformation is non-injective and implies loss of information. This chapter shows that the choice of the coordinate system to reference points in the 3D space is important. By selecting a suitable reference frame, the intrinsic nature of image formation process is kept unchanged, but the mathematical relationship between the world and the image becomes simpler and more intuitive. This can help not only the understanding of the imaging process but also the development of new algorithms and applications.

A general framework to select the most suitable coordinate system for a certain sensor/system is presented. Two differential constraints are defined to enable the choice of a 3D reference frame: the compactness constraint and the decoupling constraint. It is shown that coordinate transformations satisfying these differential constraints bring advantageous properties when mapping 3D space velocities into 2D image velocities. The derived framework is applied to conventional perspective cameras and then generalized to central catadioptric ones. The advantageous of using this approach for active tracking applications are discussed in the second part of the chapter. Three different cases are considered: a perspective camera with translational motion in the XY plane, a perspective camera with rotational pan and tilt motion and a parabolic omnidirectional camera with a rotational degree of freedom around the Z axis.

Chapter 7

Pose Estimation Using Central Panoramic Imaging

Visual servoing can benefit from sensors providing large fields of view. The previous chapter proposes an application where an omnidirectional camera is used for active tracking of a moving target. The visual information is used to control the rotation of the MVS platform. Since the robot has less than 3 degrees of freedom (DOF), the target is modeled as a point moving in the scene. However there are several applications where the target must be modeled as a rigid body with translation and rotation motion. This chapter focuses on pose estimation using central panoramic imaging.

Consider the task of positioning a robotic manipulator, with more than 3 DOF, using visual information. The approaches to this problem are traditionally classified in two groups: image based and position based visual servoing [43]. In the former the task function is defined in the image plane [28]. In the latter the control input is defined in the 3D task space. The pose of the target is estimated from image features based on the knowledge of a geometric model of the object and the camera calibration [72]. With only one camera there are ambiguities and singularities in pose estimation and the target can get out of the field of view during the tracking. In [48] a multiple camera approach is used to cope with these difficulties. Panoramic imaging can overcome the problems avoiding multiple view geometry and calibration of several cameras.

This chapter introduces the jacobian matrix \mathbf{J} for a generic central catadiop-

7. Pose Estimation Using Central Panoramic Imaging

tric system. Matrix \mathbf{J} is derived from the central catadioptric mapping function presented in chapter 2. According to this unifying theory, central catadioptric imaging can be modeled by a non linear function \tilde{h} , with the type of sensor and shape of the mirror described by a parameter ξ (Tab 2.3). For the particular case of a conventional perspective camera the parameter ξ is null. Thus, by assuming $\xi = 0$, the general central projection jacobian matrix \mathbf{J}_g becomes the well known interaction matrix \mathbf{J}_p [28]. Moreover it is shown that the derived jacobian matrix can be decomposed in the product of two matrices \mathbf{J}_c and \mathbf{J}_p ($\mathbf{J}_g = \mathbf{J}_c \cdot \mathbf{J}_p$). \mathbf{J}_c is a 2×2 matrix that is always invertible which proves that the general catadioptric jacobian \mathbf{J}_g has exactly the same singularities as the standard perspective jacobian \mathbf{J}_p [23, 49].

Experiments on iterative pose estimation from points in the catadioptric image are performed. The singularities of \mathbf{J}_g and the stability and convergence of image based visual servoing from catadioptric images are discussed. Point-to-contour tracking [48] on omnidirectional images is used to estimate the rigid displacement of objects. The application of the derived framework to control the position of a robotic arm is also discussed.

7.1 Problem Formulation

Fig. 7.1 depicts a moving rigid object observed by a central catadioptric sensor. The referential frame \mathcal{R}_b is attached to the moving body. The coordinates system of the panoramic sensor is \mathcal{R} which is centered in the effective viewpoint. The 3×3 matrix \mathbf{R} provides the rotation between \mathcal{R}_b and \mathcal{R} . The 3×1 vector $\mathbf{t} = (t_x, t_y, t_z)$ is the translation vector corresponding to the position of the origin \mathbf{O}_b in sensor coordinates. Our goal is to estimate the pose of the rigid body knowing the coordinates $\{\mathbf{X}_b^1, \mathbf{X}_b^2, \dots, \mathbf{X}_b^N\}$ of a set of N object points.

7.1.1 The Central Panoramic Sensor

Chapter 2 presents an unifying theory for central catadioptric image formation. The general mapping model is represented in Fig. 2.3. If the system is calibrated then collineation \mathbf{H}_c , given in equation 2.7, is known. For the purposes of the present chapter it is assumed, without loss of generality, that $\mathbf{H}_c = \mathbf{I}$ with \mathbf{I} the

7.1. Problem Formulation

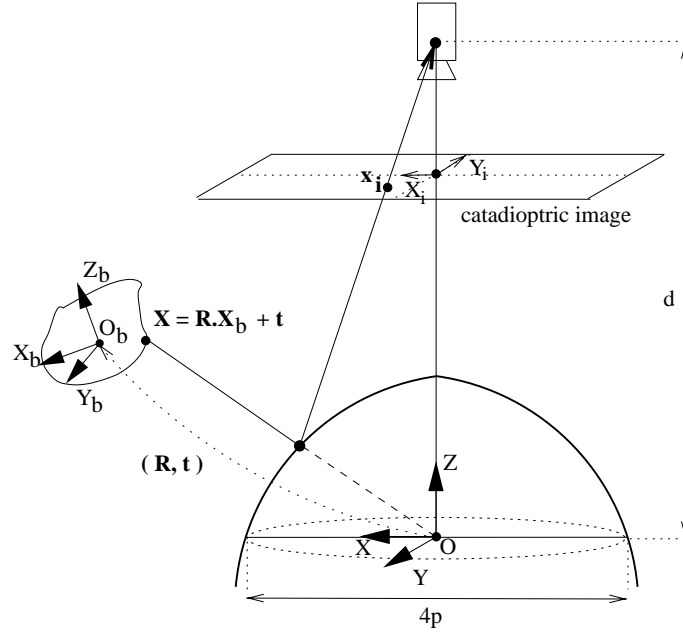


Figure 7.1: Central catadioptric projection of a rigid body

3×3 identity matrix.

Assume a generic 3D point in the scene, with sensor coordinates $\mathbf{X} = (X, Y, Z)^t$, which is projected at point $\mathbf{x}_i = (x_i, y_i)^t$ in the catadioptric image plane (Fig. 7.1). Consider the scheme of Fig. 2.3 where $\mathbf{X}_h = (\mathbf{X}^t, 1)^t$. Since $\mathbf{P} = [\mathbf{I}|\mathbf{0}]$ it comes in a straightforward manner that $\mathbf{x} = (X, Y, Z)^t$. The oriented projective ray \mathbf{x} is mapped into point $\bar{\mathbf{x}}$ by the non-linear function \bar{h} (equation 2.5). Collineation \mathbf{H}_c transforms the projective point $\bar{\mathbf{x}}$ into point $\hat{\mathbf{x}}$ (equation 2.6). Taking into account that $\mathbf{H}_c = \mathbf{I}$ yields

$$\hat{\mathbf{x}} = \left(\frac{X}{\sqrt{X^2 + Y^2 + Z^2}}, -\frac{Y}{\sqrt{X^2 + Y^2 + Z^2}}, \frac{Z}{\sqrt{X^2 + Y^2 + Z^2}} + \xi \right)^t$$

The projective ray $\hat{\mathbf{x}} = (\hat{x}, \hat{y}, \hat{z})^t$ intersects the catadioptric image plane on point \mathbf{x}_i with coordinates $(x_i, y_i)^t = \left(\frac{\hat{x}}{\hat{z}}, \frac{\hat{y}}{\hat{z}} \right)^t$. Function \mathbf{f}_i , provided in equation 7.2, maps 3D scene points into 2D image points such that $\mathbf{x}_i = \mathbf{f}_i(\mathbf{X})$.

7. Pose Estimation Using Central Panoramic Imaging

$$\mathbf{f}_i() : (X, Y, Z) \longrightarrow \left(\frac{X}{Z + \xi\sqrt{X^2 + Y^2 + Z^2}}, -\frac{Y}{Z + \xi\sqrt{X^2 + Y^2 + Z^2}} \right) \quad (7.2)$$

7.1.2 Measuring the Pose Estimation Error

Consider the scheme depicted in Fig. 7.1 where \mathbf{X}_b are the body coordinates of a generic point lying on the moving object. By body coordinates we mean coordinates in the reference frame attached to the rigid object. The coordinate system \mathfrak{R} attached to the panoramic camera is called the sensor reference frame. If the pose (\mathbf{R}, \mathbf{t}) between the two coordinates systems is known then the point 3D position in sensor coordinates is

$$\mathbf{X} = \mathbf{R}\mathbf{X}_b + \mathbf{t} \quad (7.3)$$

Assume that the object moves rigidly with relation to the catadioptric sensor. The change in pose can be described in a differential way by a 6×1 kinematic screw $\delta = (\boldsymbol{\omega}, \mathbf{v})^t$ [51]. Consider the point $\mathbf{X} = (X, Y, Z)^t$ lying on the object. The 3D velocity of the point in sensor coordinates is

$$\dot{\mathbf{X}} = \underbrace{\begin{bmatrix} 0 & -Z & Y & 1 & 0 & 0 \\ Z & 0 & -X & 0 & 1 & 0 \\ -Y & X & 0 & 0 & 0 & 1 \end{bmatrix}}_{\mathbf{J}_m} \delta \quad (7.4)$$

Point \mathbf{X} is projected into the catadioptric image plane at point $\mathbf{x}_i = \mathbf{f}_i(\mathbf{X})$ with \mathbf{f}_i the mapping function provided in equation 7.2. If \mathbf{J}_i is the 2×3 jacobian matrix of function \mathbf{f}_i then the velocity of the image point due to the object rigid motion is provided by

$$\dot{\mathbf{x}}_i = \underbrace{\mathbf{J}_i \mathbf{J}_m}_{\mathbf{J}_g} \delta \quad (7.5)$$

We aim to estimate the pose of the rigid body from its catadioptric image. Consider a set $\{\mathbf{X}_b^1, \mathbf{X}_b^2, \dots, \mathbf{X}_b^N\}$ of 3D points lying on the object. The position of these points in body coordinates is known in advance. This set of points will be used as a 3D model of the rigid object. If the rotation matrix is \mathbf{R} and

7.1. Problem Formulation

the translation vector is \mathbf{t} then the set of 3D points is imaged in a set of points $\mathbf{s} = \{\mathbf{x}_i^1, \mathbf{x}_i^2, \dots, \mathbf{x}_i^N\}$. The image points can be determined using the relation of equation 7.6.

$$\mathbf{x}_i^j = \mathbf{f}_i(\mathbf{R} \cdot \mathbf{X}_b^j + \mathbf{t}), \quad j=1,2,3 \dots N \quad (7.6)$$

Assume an initial pose estimate $(\check{\mathbf{R}}, \check{\mathbf{T}})$. Let $\check{\mathbf{s}} = \{\check{\mathbf{x}}_i^1, \check{\mathbf{x}}_i^2, \dots, \check{\mathbf{x}}_i^N\}$ be the corresponding catadioptric projection of the set of N model points. If the pose estimation is correct then $\check{\mathbf{s}}$ is coincident with the set of real image points \mathbf{s} . However in general there is an error in the estimate. Vector \mathbf{e} is defined as $\mathbf{e} = \mathbf{s} - \check{\mathbf{s}}$ and depends on the pose estimate error.

The error in the pose estimation can be described by a kinematic screw δ [51, 28, 23]. From the result of equation 7.5 comes that $(\mathbf{x}_i^j - \check{\mathbf{x}}_i^j) \approx \mathbf{J}_g^j \delta$ with $j = 1, 2 \dots N$ and \mathbf{J}_g^j the jacobian matrix \mathbf{J}_g evaluated on the j^{th} model point. Equation 7.7 establishes the relationship between the measured image error \mathbf{e} and the error δ on the pose estimation of the rigid body. \mathbf{J} is a $2N \times 6$ matrix comprised by the jacobian matrix \mathbf{J}_g evaluated in the N points of the object model.

$$\underbrace{\begin{bmatrix} \mathbf{x}_i^1 - \check{\mathbf{x}}_i^1 \\ \mathbf{x}_i^2 - \check{\mathbf{x}}_i^2 \\ \vdots \\ \mathbf{x}_i^N - \check{\mathbf{x}}_i^N \end{bmatrix}}_{\mathbf{e}} = \underbrace{\begin{bmatrix} \mathbf{J}_g^1 \\ \mathbf{J}_g^2 \\ \vdots \\ \mathbf{J}_g^N \end{bmatrix}}_{\mathbf{J}} \delta \quad (7.7)$$

Given the initial pose estimate $(\check{\mathbf{R}}, \check{\mathbf{T}})$, the points of the 3D model of the object are projected in the set of image points $\check{\mathbf{s}}$ using the relation provided in equation 7.6. The error vector \mathbf{e} is the difference between points in \mathbf{s} and $\check{\mathbf{s}}$. Vector \mathbf{e} is measured in the image plane and corresponds to the distance between the model projection and the real image of the object. According to the result of equation 7.7 the image error \mathbf{e} is related to the pose estimation error δ by the jacobian matrix \mathbf{J} . Thus knowing both \mathbf{e} and \mathbf{J} the pose error can be determined using normal least squares [58].

$$\delta = (\mathbf{J}^t \mathbf{J})^{-1} \mathbf{J}^t \mathbf{e} \quad (7.8)$$

The objective is to update the pose estimation such that the projection of the

7. Pose Estimation Using Central Panoramic Imaging

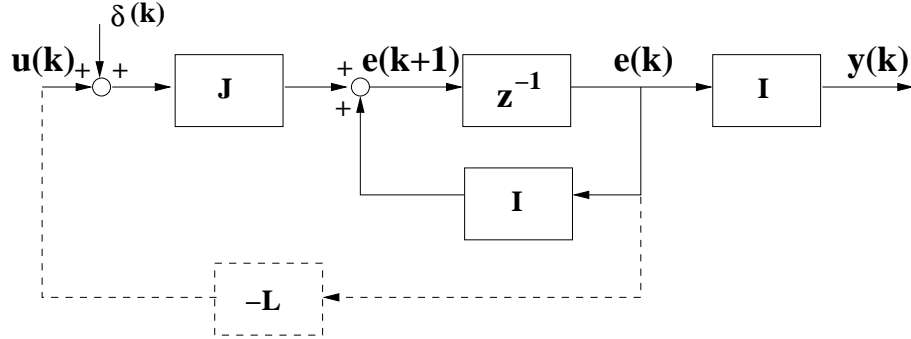


Figure 7.2: Iterative pose estimation as a regulation control problem (I is the $2N \times 2N$ identity matrix).

3D model becomes coincident with the object image and the measured error vector e converges to zero. This problem is known in the literature as model based tracking of a rigid object.

7.1.3 Pose Estimation as a Regulation Control Problem

The problem of model based tracking stated in the previous section can be formulated as a regulation control problem [3]. Consider the system depicted in Fig. 7.2 with input u and output y . The corresponding state space model is provided in equation 7.9. The state vector is the image error e and the input matrix is the jacobian J (equation 7.7).

$$\begin{cases} e(k+1) = e(k) + J(u(k) + \delta(k)) \\ y(k) = e(k) \end{cases} \quad (7.9)$$

Accordingly to the system state-space equations the change in pose δ acts as an input perturbation which disturbs the output $y(k)$. The goal is to find a state feedback controller L such that if $u(k) = -Le(k)$ then the disturbance is rejected and the system state vector e converges to zero. Or, in other words, the projection of the 3D model and the image of the object become coincident.

Consider the result of equation 7.8 where the pose estimation error δ is determined using the image error e . System regulation can be achieved by making $u = \delta$ (equation 7.10). Equation 7.11 provides the state space model of the fi-

7.2. The Jacobian Matrix for General Central Catadioptric Projection

nal closed loop system. System stability and transient response depend on the eigenvalues of the matrix $(\mathbf{I} - \mathbf{J}(\mathbf{J}^t\mathbf{J})^{-1}\mathbf{J}^t)$ [3]. However it is important to remind that the state transition matrix is a function of rigid body position which changes along time. Thus the analysis of the regulation dynamics is hard to perform. Moreover the controller of equation 7.10 is only realizable when $(\mathbf{J}^t\mathbf{J})$ is non singular. Whenever matrix $(\mathbf{J}^t\mathbf{J})$ is not invertible we are in presence of a singularity [49, 23].

$$\mathbf{L} = (\mathbf{J}^t\mathbf{J})^{-1}\mathbf{J}^t \quad (7.10)$$

$$\begin{cases} \mathbf{e}(\mathbf{k} + 1) = (\mathbf{I} - \mathbf{J}(\mathbf{J}^t\mathbf{J})^{-1}\mathbf{J}^t)\mathbf{e}(\mathbf{k}) + \mathbf{J}\delta(\mathbf{k}) \\ \mathbf{y}(\mathbf{k}) = \mathbf{e}(\mathbf{k}) \end{cases} \quad (7.11)$$

7.2 The Jacobian Matrix for General Central Catadioptric Projection

The design the controller of equation 7.10 requires the derivation of matrix \mathbf{J} . Matrix \mathbf{J} depends on the jacobian matrix \mathbf{J}_g which is evaluated on the N points of the object model (equation 7.7). This section derives matrix \mathbf{J}_g .

7.2.1 The Jacobian Matrix \mathbf{J}_g

Consider the central catadioptric mapping function \mathbf{f}_i which maps 3D point coordinates \mathbf{X} in image coordinates \mathbf{x}_i . The corresponding jacobian matrix \mathbf{J}_i is derived by differentiating the function of equation 7.2. The obtained result is presented in equation 7.12 where $\rho = \sqrt{X^2 + Y^2 + Z^2}$.

$$\mathbf{J}_i = \frac{1}{\rho(Z + \xi\rho)^2} \begin{bmatrix} \rho Z + \xi(Y^2 + Z^2) & -\xi XY & -X(\rho + \xi Z) \\ \xi XY & -(\rho Z + \xi(X^2 + Z^2)) & Y(\rho + \xi Z) \end{bmatrix} \quad (7.12)$$

Consider the 3×3 matrix \mathbf{J}_m provided in equation 7.4. Notice that $\mathbf{J}_m = [\tilde{\mathbf{X}}|\mathbf{I}]$ with $\tilde{\mathbf{X}}$ the skew symmetric matrix associated with point coordinates \mathbf{X} (equation 4.3). From equation 7.5 results that $\mathbf{J}_g = \mathbf{J}_i\mathbf{J}_m$. The jacobian matrix \mathbf{J}_g for

7. Pose Estimation Using Central Panoramic Imaging

general central catadioptric projection is provided in equation 7.13.

$$\mathbf{J}_g = \begin{bmatrix} \frac{XY}{(Z+\xi\rho)^2} & -\frac{X^2+Z^2+\xi\rho Z}{(Z+\xi\rho)^2} & -\frac{Y}{(Z+\xi\rho)} & \frac{\rho Z+\xi(Y^2+Z^2)}{\rho(Z+\xi\rho)^2} \\ -\frac{Y^2+Z^2+\xi\rho Z}{(Z+\xi\rho)^2} & \frac{XY}{(Z+\xi\rho)^2} & \frac{X}{(Z+\xi\rho)} & \frac{\xi XY}{\rho(Z+\xi\rho)^2} \\ & & & -\frac{\xi XY}{\rho(Z+\xi\rho)^2} & -\frac{X(\rho+\xi Z)}{\rho(Z+\xi\rho)^2} \\ & & & -\frac{\rho Z+\xi(X^2+Y^2)}{\rho(Z+\xi\rho)^2} & \frac{Y(\rho+\xi Z)}{\rho(Z+\xi\rho)^2} \end{bmatrix} \quad (7.13)$$

Equation 7.13 presents the jacobian matrix \mathbf{J}_g as a function of the 3D coordinates of point $\mathbf{X} = (X, Y, Z)^t$. Function Γ , provided in Tab. 6.3 (chapter 6), transforms the cartesian coordinates $\mathbf{X} = (X, Y, Z)^t$ in the special coordinates $\Upsilon = (\phi, \psi, \rho)^t$ (Fig. 6.4). Replacing in equation 7.13 (X, Y, Z) by $\mathbf{T}(\Upsilon)$, also provided in Tab. 6.3, matrix \mathbf{J}_g is written in terms of (ϕ, ψ, ρ) . Taking into account that $\phi = \arctan(x_i)$ and $\psi = \arctan(y_i)$ yields

$$\mathbf{J}_g = \begin{bmatrix} -x_i y_i & -\frac{(1+x_i^2)\Xi - y_i^2 \xi}{\Xi + \xi} & -y_i & \frac{1+x_i^2(1-\xi(\Xi+\xi))+y_i^2}{\rho(\Xi+\xi)} \\ -\frac{(1+y_i^2)\Xi - x_i^2 \xi}{\Xi + \xi} & -x_i y_i & x_i & -\frac{x_i y_i \xi}{\rho} \\ & & & \frac{x_i y_i \xi}{\rho} & -\frac{x_i \Xi}{\rho} \\ & & & -\frac{1+x_i^2+y_i^2(1-\xi(\Xi+\xi))}{\rho(\Xi+\xi)} & -\frac{y_i \Xi}{\rho} \end{bmatrix} \quad (7.14)$$

Equation 7.14 provides the central catadioptric jacobian matrix \mathbf{J}_g as a function the image coordinates x_i and the point depth ρ ($\Xi = \sqrt{1 + (x_i^2 + y_i^2)(1 - \xi^2)}$). This representation is much more useful from an application point of view.

7.2.2 Additional Considerations

Making $\xi = 0$ in equation 7.13 yields

$$\mathbf{J}_p = \begin{bmatrix} \frac{XY}{Z^2} & -(1 + \frac{X^2}{Z^2}) & -Y & \frac{1}{Z} & 0 & -\frac{X}{Z^2} \\ -(1 + \frac{Y^2}{Z^2}) & \frac{XY}{Z^2} & X & 0 & -\frac{1}{Z} & \frac{Y}{Z^2} \end{bmatrix} \quad (7.15)$$

Matrix \mathbf{J}_p is the well known jacobian matrix for conventional perspective cameras. Matrix \mathbf{J}_g , derived above, is a generalization of the interaction matrix intro-

7.2. The Jacobian Matrix for General Central Catadioptric Projection

duced in [28].

The jacobian \mathbf{J}_i , provided in equation 7.12, can be decomposed in the matrix product of equation 7.16. The 2×2 matrix \mathbf{J}_c depends on point coordinates \mathbf{X} and on the mirror parameter ξ . Notice that if $\xi = 0$ then matrix \mathbf{J}_c becomes the 2×2 identity matrix. The second matrix has dimension 2×3 and is the jacobian of the conventional perspective mapping function $\mathbf{f}_i = (X/Z, -Y/Z)^t$ obtained making ξ null in equation 7.2. The eigenvalues of matrix \mathbf{J}_c are $\{Z/(Z + \rho\xi); (Z^2(\rho + \xi Z))/(\rho(Z + \xi\rho)^2)\}$. Notice that \mathbf{J}_c is positive definite whenever the Z coordinate is positive.

$$\mathbf{J}_i = \underbrace{\begin{bmatrix} \frac{Z(\rho Z + \xi(Y^2 + Z^2))}{\rho(Z + \xi\rho)^2} & \frac{\xi XYZ}{\rho(Z + \xi\rho)^2} \\ \frac{\xi XYZ}{\rho(Z + \xi\rho)^2} & \frac{Z(\rho Z + \xi(X^2 + Z^2))}{\rho(Z + \xi\rho)^2} \end{bmatrix}}_{\mathbf{J}_c} \begin{bmatrix} \frac{1}{Z} & 0 & -\frac{X}{Z^2} \\ 0 & -\frac{1}{Z} & \frac{Y}{Z^2} \end{bmatrix} \quad (7.16)$$

From the result of equation 7.16 arises that the general catadioptric matrix \mathbf{J}_g can be written as $\mathbf{J}_g = \mathbf{J}_c \mathbf{J}_p$. Matrix \mathbf{J}_p is the 2×6 jacobian matrix for the conventional perspective camera situation (equation 7.15).

The controller of equation 7.10 is realizable if, and only if, \mathbf{J} is a full rank matrix. \mathbf{J} has dimension $2N \times 6$ with N the number of points in the 3D model of the object. Clearly the full rank constraint can not be verified with less than three points. Equation 7.17 is derived from equation 7.7 knowing that $\mathbf{J}_g^j = \mathbf{J}_c^j \mathbf{J}_p^j$.

$$\mathbf{J} = \underbrace{\begin{bmatrix} \mathbf{J}_c^1 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{J}_c^2 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{J}_c^N \end{bmatrix}}_{\mathbf{C}} \underbrace{\begin{bmatrix} \mathbf{J}_p^1 \\ \mathbf{J}_p^2 \\ \vdots \\ \mathbf{J}_p^N \end{bmatrix}}_{\mathbf{P}} \quad (7.17)$$

Matrix \mathbf{J} is the product of a $2N \times 2N$ square matrix \mathbf{C} with a matrix \mathbf{P} with dimension $2N \times 6$. Assuming $Z > 0$ then all matrices \mathbf{J}_c^j are positive definite and matrix \mathbf{C} is always full rank. Matrix \mathbf{J} is rank deficient if, and only if, matrix \mathbf{P} is also rank deficient. This proves that the general central catadioptric situation presents the same singularities of the perspective case. These singularities have been studied in detail in [49, 48].

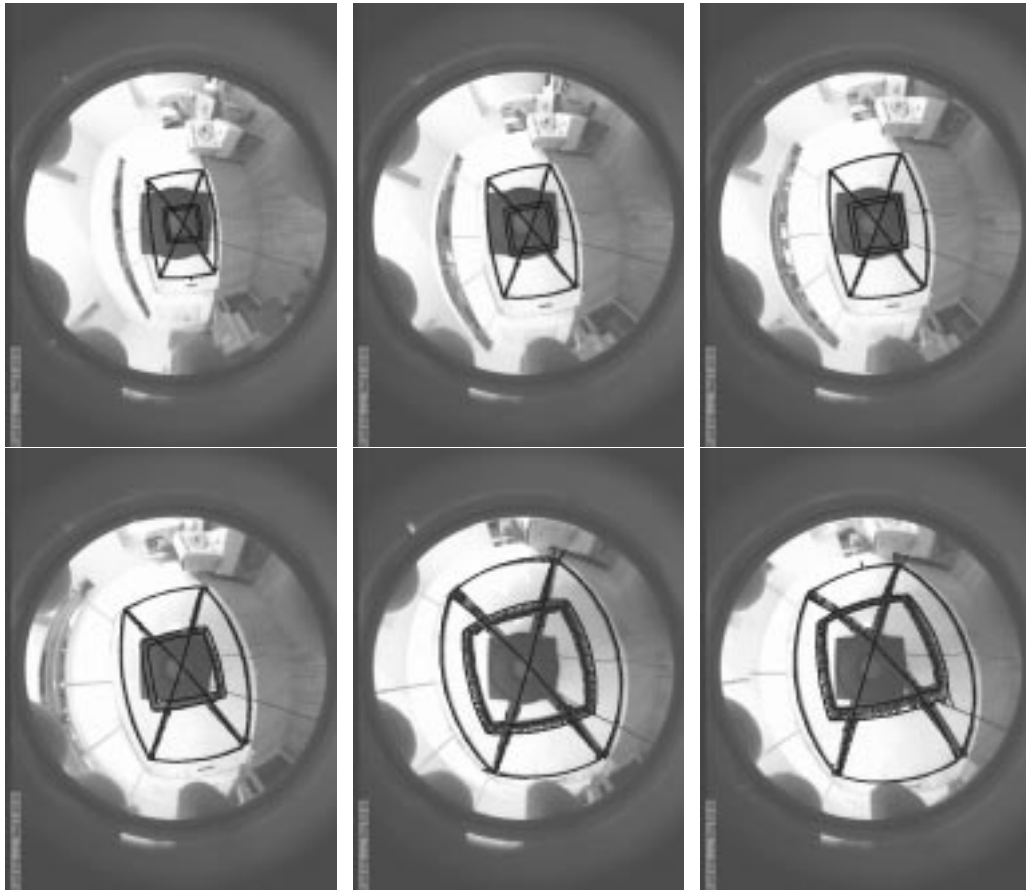


Figure 7.3: A tracking sequence. The object translates along axis of camera

7.3 Model Based Tracking

Based on the tracking method described above we implemented an object tracker. Since with catadioptric cameras straight lines map onto the image plane as conics, we devised a contour-to-point tracker along the lines described in [48]. The figures below show a rectangular object moving towards the camera and in a direction perpendicular to the camera.

7.4 Closure

This chapter introduces for the first time the jacobian matrix \mathbf{J}_g for a general central catadioptric system. It is shown that the conventional interaction matrix \mathbf{J}_p

7.4. Closure

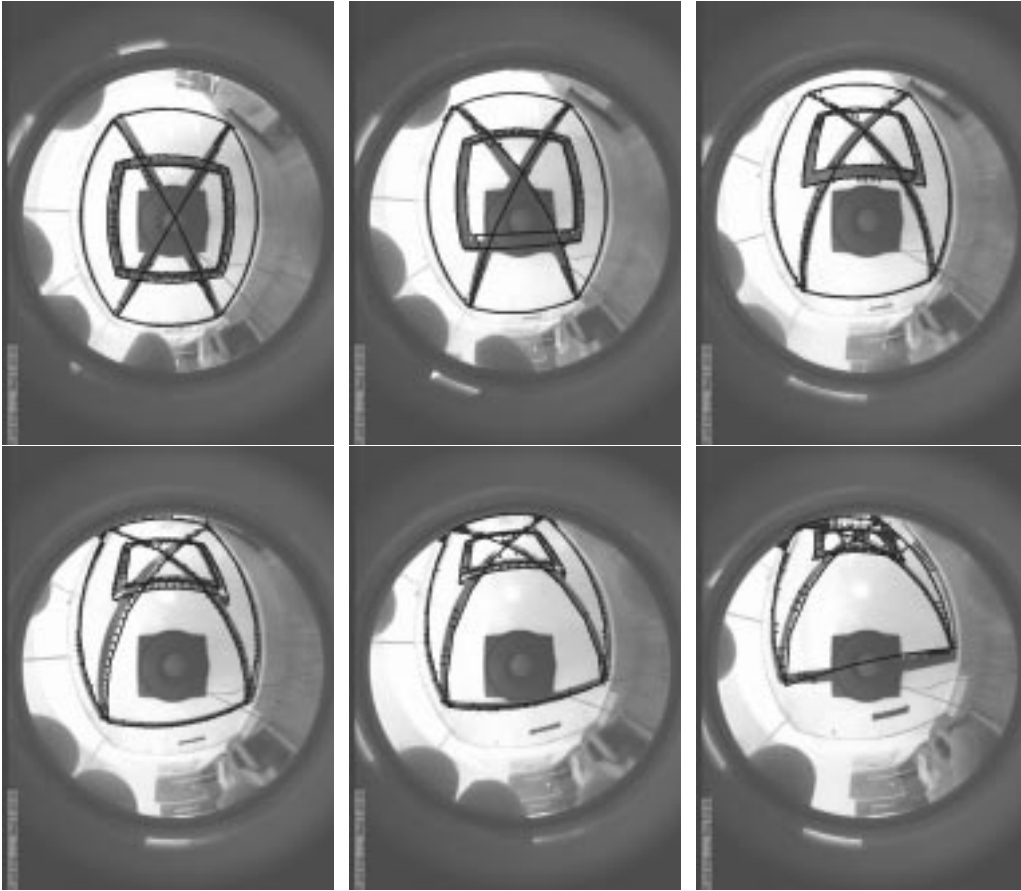


Figure 7.4: A tracking sequence. The object translates in front of camera..

for the perspective cameras [28] is a particular case of matrix \mathbf{J}_g . Moreover the general jacobian matrix \mathbf{J}_g can be written as the product of matrices \mathbf{J}_c and \mathbf{J}_p ($\mathbf{J}_g = \mathbf{J}_c \cdot \mathbf{J}_p$). Since \mathbf{J}_c is a 2×2 non singular matrix encoding the mirror information, the general catadioptric jacobian \mathbf{J}_g has exactly the same singularities as the standard perspective jacobian \mathbf{J}_p [23, 49].

The general jacobian matrix can be used to extend to central panoramic imaging algorithms and techniques originally developed for perspective cameras. As an example in [71] Vidal et al. propose a factorization approach for motion segmentation and 3D motion estimation from paracatadioptric images which requires the jacobian matrix. In this chapter we present some experiments in estimating the pose of a rigid body imaged by a central catadioptric sensor. Given an initial pose estimate, the error between the projected model and the actual image of the object

7. Pose Estimation Using Central Panoramic Imaging

is measured. The pose parameters are updated by minimizing the image error. Since the objective function is non-linear, the estimation must be performed using a gradient descending method which requires the derived jacobian matrix. Instead of studying the pose estimation in a conventional optimization framework we formulated the problem in terms of model-based tracking. The model based tracking can be interpreted as a regulation control problem in the image plane. The motion of the rigid body acts as a disturbance that must be compensated.

The model based tracking of a rigid object can be exploited in many ways for visual servoing applications. The proposed approach is being used in robot navigation and cooperation [56]. The experimental setup consists in two mobile platforms both equipped with central catadioptric cameras. A visual landmark, similar to the one depicted in the figures, is positioned on the room ceiling. One robot is the leader with independent motion and the other is the slave. The objective is to control slave motion such that the relative position between the two platforms is kept constant. To achieve this goal both robots use the omnidirectional vision to estimate their pose from the model based tracking of the landmark. A method to control the position of a robotic arm using a static catadioptric system is also being developed. Typically, in visual servoing using a conventional perspective camera, the available field of view is only enough to image the region around the end-effector. The pose of the end-effector is estimated by visual feedback, and motion control is achieved using the manipulator jacobian known “a priori”. The success of this approach is highly dependent on the arm calibration. We use the wide field of view provided by the omnidirectional sensor to image the entire arm. The different manipulator links are tracked in the catadioptric image and the motion of each joint is estimated. This approach increases the robustness and accuracy of the visual servoing.

Chapter 8

Final Remarks

This work starts by introducing an image formation model for central catadioptric systems. The model covers the conventional perspective camera, the paracatadioptric sensor and systems combining a perspective camera with an hyperbolic, elliptical or planar mirrors as described in [5]. In [53] Nayar and Veri prove that any folded system that uses conic mirrors has a geometrically equivalent system using a single conic mirror. Thus the folded catadioptric camera with a single projection center can also be modelled using the proposed framework. Another way to obtain central panoramic imaging is by using cameras equipped with fish-eye lens [6]. It is not clear that our framework can be used to model these systems as well. Such an extension would be an interesting direction of research.

Chapter 3 studies in detail central catadioptric line projection. Several invariant properties are derived that are useful both for calibration and reconstruction. It is proved that any catadioptric system can be fully calibrated from the image of three lines, and that only two lines are required when the mirror is hyperbolic/elliptical and the pose between the camera and the mirror is known in advance. It would be interesting to investigate the projection of quadric surfaces by central catadioptric systems. Preliminary studies show that the image of a sphere is a conic curve, but more research must be done on this topic.

General purpose algorithms to work with conic curves are proposed on chapter 4. A closed form formula that computes the intersection between a line and a conic is derived. A method to determine the intersections between two conics is presented as well. The intersection points are computed by solving a third or-

8. Final Remarks

der equation in a single variable. The implementation of the algorithm is simple and numerically stable. We also review the most common used conic fitting techniques. It is shown that the estimation results are poor in the presence of strong occlusion. Since in general only a small arc of the line image is visible in the catadioptric image plane, then the accurate estimation of the corresponding conic locus using standard techniques is hard to accomplish.

In chapter 5 we overcome the problem of estimating lines images for the case of parabolic systems. We derive for the first time the necessary and sufficient conditions that must be verified by a set of conics to be the paracatadioptric projection of a set of lines. This allows us to constrain the search space and improve the performance of the conic fitting. A robust method to calibrate the sensor using line images is proposed. Experimental results show that it works significantly better than any other algorithm presented so far [74, 45, 35]. Additionally we propose a computationally efficient algorithm to estimate lines in the paracatadioptric image plane. The method uses the necessary and sufficient conditions for a conic to be the paracatadioptric projection of a line. Prior knowledge of system calibration is required. Experimental results show that the approach is very robust and the estimation results are much better than the ones obtained by performing perspective rectification. We intend to pursue the research to obtain similar algorithms for hyperbolic/elliptical sensors. The major difficulty is that the necessary and sufficient conditions for a conic to be the image of a line are highly non-linear.

Chapter 6 dicusses the importance of chosing a suitable world coordinate system when modelling a general central projection sensor. This is an original topic which, as far as we know, has never been mentioned in the literature. The choice of a certain reference frame does not imply that new information will be available in the images. Instead the geometric transformations can be represented in a common and more compact way. A generic framework to select the most suitable world coordinate system for a certain sensor/application is introduced. Examples of active vision systems which benefit from an adequate choice of the reference frame are provided. The jacobian matrix for general central projection systems is studied in chapter 7. This matrix is a generalization of the conventional interaction matrix for perspective cameras [28]. The derived results are applied to perform model based tracking of a rigid body on the catadioptric image plane.

An important topic that is not covered at all in the present thesis is the mul-

multiple view geometry of general central catadioptric systems. The first work that we are aware of on the epipolar geometry of general central projection systems has been presented by Svoboda et al. [68]. Since then other authors have developed efforts to derive a general fundamental matrix. In [36, 67] the 2D image plane is embedded in a higher dimensional space as a way to linearize the relations for the paracatadioptric image formation. Geyer and Daniilidis come up with a 4×4 fundamental matrix for parabolic systems [36]. Nevertheless this result requires the system to be skewless with unitary aspect ratio. In [67] Sturm derives a more general fundamental relation which mixes parabolic systems with conventional perspective cameras. More recently Micusik and Padjla explore the distortion model proposed in [31] to calibrate a camera with fish-eye lens from epipolar geometry [50]. Despite of these works, the derivation of a fundamental matrix for general central projection systems which generalize the well known results for perspective cameras is still an open problem.

8. Final Remarks

Bibliography

- [1] Daniel G. Aliaga. Accurate calibration for real-time pose estimation in room-size environment. In *in Proc. of the IEEE International Conference on Computer Vision*, Vancouver, Canada, July 2001.
- [2] Y. Aloimonos, I. Weiss, and A. Bandyopadhyay. Active vision. *International Journal of Computer Vision*, 1(4):333–356, January 1988.
- [3] K. Astrom and B. Wittenmark. *Computer Controlled Systems. Theory and Design*. Prentice Hall, 1997.
- [4] P. Baker, C. Fermuller, Y. Aloimonos, and R. Pless. A spherical eye from multiple cameras (makes better models of the world. In *in Proc. of the IEEE Int. Conf. on Computer Vision and Pattern Recognition*, Kauai,Haway,USA, December 2001.
- [5] S. Baker and S. Nayar. A theory of catadioptric image formation. In *in Proc. of IEEE International Conference on Computer Vision*, pages 35 – 42, Bombay, 1998.
- [6] H. Bakstein and T. Padjla. Non-central cameras: A review. In *in Proc. of Computer Vision Winter Workshop*, pages 223 – 233, Ljubljana, Slovenia, 2001.
- [7] Joao P. Barreto and Helder Araujo. Issues on the geometry of central catadioptric imaging. In *in Proc. of the IEEE Int. Conf. on Computer Vision and Pattern Recognition*, Kauai, Haway, USA, December 2001.

BIBLIOGRAPHY

- [8] Joao P. Barreto and Helder Araujo. Geometric properties of central catadioptric line images. In *in Proc. of European Conference on Computer Vision*, Copenhagen, Denmark, May 2002.
- [9] Joao P. Barreto and Helder Araujo. Direct least square fitting of paracatadioptric line images. In *Proc. of OMNIVIS'2003 - Workshop on Omnidirectional Vision and Camera Networks*, Madison, Wisconsin, June 2003.
- [10] Joao P. Barreto and Helder Araujo. A general framework for selecting world coordinate systems in perspective and catadioptric imaging application. *International Journal on Computer Vision (to appear)*, 2003.
- [11] Joao P. Barreto and Helder Araujo. Paracatadioptric camera calibration using lines. In *Proc. of ICCV'2003 - IEEE Int. Conf. on Computer Vision*, Nice, France, October 2003.
- [12] Joao P. Barreto, Frederick Martin, and Radu Horaud. Visual servoing/tracking using central catadioptric images. In *in ISER2002 - Proc. of 8th Int. Symposium on Experimental Robotics*, Ischia, Italy, July 2002.
- [13] João P. Barreto, J. Batista, and H. Araújo. Tracking multiple targets in 3d. In *IROS99-IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 233–246, Kyongju, Korea, October 1999.
- [14] João P. Barreto, P. Peixoto, J. Batista, and H. Araújo. Improving 3d active visual tracking. In Henrik I. Christensen, editor, *Computer Vision Systems, Lectures Notes in Computer Science 1542*. Springer, 1999.
- [15] João P. Barreto, P. Peixoto, J. Batista, and H. Araujo. Evaluation of the robustness of visual behaviors through performance characterization. In Markus Vincze and Gregory D. Hager, editors, *Robust Vision for Vision-Based Control of Motion*. IEEE Press, 1999.
- [16] J. Batista, J. Dias, H. Araújo, and A. Almeida. The isr multi-degree-of-freedom active vision robot head: Design and calibration. In *in Proc. of M2VIP'95 - Second International Conference on Mechatronics and Machine Vision in Practice*, Hong-Kong, September 1995.

BIBLIOGRAPHY

- [17] J. Batista, P. Peixoto, and H. Araujo. Real time visual behavior with a binocular vision system. In *Proc. of the IEEE Int. Conf. on Robotics and Automation*, New Mexico, USA, September 1997.
- [18] R. Benosman and S. Kang. *Panoramic Vision*. Springer Verlag, 2000.
- [19] F. Bookstein. Fitting conic sections to scattered data. *Computer Vision, Graphics and Image Processing*, 9, 1979.
- [20] T. E. Boulton, R. J. Micheals, M. Eckmann, X. Gao, C. Power, and S. Sablak. Omnidirectional video applications. In *Proc. of the 8th Int. Symposium on Intelligent Robotic Systems*, Reading, UK, July 2000.
- [21] J. F. Canny. A computational approach to edge detection. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 8(6), June 1986.
- [22] J. Chahl and M. Srinivasan. Reflective surface for panoramic imaging. *Applied Optics*, 36(31):8275–8285, 1997.
- [23] Francois Chaumette. *Potential Problems of Stability and Convergence in Image Based and Position Based Visual Servoing*, volume 237 of *Lecture Notes in Control and Information Systems*. Springer-Verlag, 1998.
- [24] S. Chroust, M. Vincze, R. Traxl, and P. Krautgartner. Evaluation of processing architecture and control law on the performance of vision based control systems. In *AMC00–IEEE/RSJ Int. Workshop on Advanced Motion Control*, pages 19–24, Japan, April 2000.
- [25] Peter I. Corke. *Visual Control of Robots: High Performance Visual Servoing*. Mechatronics. John Wiley, 1996.
- [26] R. Courant. *Differential and Integral Calculus*. Wiley Classic Library, 1998.
- [27] Dennis and Schmabel. *Numerical Methods for Unconstrained Optimization and Non-Linear Equations*. Classics in Applied Mathematics.
- [28] B. Espiau, F. Chaumette, and P. Rives. A new approach to visual servoing in robotics. *IEEE Trans. on Robot. and Automat.*, 8(3):313–326, June 1992.

BIBLIOGRAPHY

- [29] Olivier Faugeras. *Three Dimensional Computer Vision. A Geometric Viewpoint*. MIT Press, 1993.
- [30] Olivier Faugeras and Quang-Tuan Luong. *The Geometry of Multiple Images*. MIT Press, 2001.
- [31] A. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. In *in Proc. of IEEE Int. Conference on Computer Vision and Pattern Recognition*, Hawai, USA, 2001.
- [32] A. Fitzgibbon and R. Fisher. A buyer's guide to conic fitting. In *in Proc. of British Machine Vision Conference*, Birmingham, England, 1995.
- [33] A. Fitzgibbon, M Pilu, and R. Fisher. Direct least square fitting of ellipses. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 21(5), May 1999.
- [34] C. Geyer and K. Daniilidis. An unifying theory for central panoramic systems and practical implications. In *in Proc. of European Conference on Computer Vision*, pages 445 – 461, Dublin, 2000.
- [35] C. Geyer and K. Daniilidis. Paracatadioptric camera calibration. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 24(4), April 2002.
- [36] C. Geyer and K. Daniilidis. Properties of the catadioptric fundamental matrix. In *in Proc. of European Conference on Computer Vision*, Copenhagen, Denmark, May 2002.
- [37] J. Gluckman and S. Nayar. Egomotion and omnidirectional cameras. In *in Proc. of IEEE International Conference on Computer Vision*, pages 999 – 1005, Bombay, 1998.
- [38] Gene H. Golub and Charles F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, 1983.
- [39] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
- [40] E. Hecht and A. Zajac. *Optics*. Addison-Wesley, 1974.

BIBLIOGRAPHY

- [41] R. A. Hicks and R. Bajcsy. Catadioptric sensor that approximate wide-angle perspective projection. In *in Proc. of the IEEE Int. Conf. on Computer Vision and Pattern Recognition*, pages 545 – 551, Hilton Head Island, USA, 2000.
- [42] R. A. Hicks and R. K. Perline. Equi-areal catadioptric sensors. In *in Proc. of IEEE Workshop on Omnidirectional Vision*, pages 13 – 18, Copenhagen, Denmark, May 2002.
- [43] S. Hutchinson, G. Hager, and P. I. Corke. A tutorial on visual servo control. *IEEE Trans. on Robotics and Automation*, 12(5):651–670, October 1996.
- [44] K. Kanatani. Statistical bias of conic fitting and renormalization. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 16(3), March 1994.
- [45] S. B. Kang. Catadioptric self calibration. In *in Proc. of the IEEE Int. Conf. on Computer Vision and Pattern Recognition*, pages 201 – 207, USA, 2000.
- [46] P. Krautgartner and M. Vincze. Performance evaluation of vision based control tasks. In *in Proc. of the IEEE Int. Conf. on Robotics and Automation*, Leuven, Belgium, May 1998.
- [47] S. Laveau and O. Faugeras. Oriented projective geometry for computer vision. In *in Proc. of European Conference on Computer Vision*, 1996.
- [48] F. Martin and R. Horaud. Multiple camera tracking of rigid objects. In *INRIA Research report n 4268*, Grenoble - France, September 2001.
- [49] H. Michel and P. Rives. Singularities in the determination of the situation of a robot effector from the perspective view of 3 points. In *Research report 1850 INRIA*, Sophia-Antipolis - France, February 1993.
- [50] B. Micusik and T. Padjla. Estimation of omnidirectional camera model from epipolar geometry. In *IEEE Int. Conference on Computer Vision and Pattern Recognition*, Madison, Wisconsin, USA, June 2003.
- [51] R. Murray, Z. Li, and S. Sastry. *A Mathematical Introduction to Robotic Manipulation*. CRC Press, 1994.

BIBLIOGRAPHY

- [52] V. S. Nalwa. A true omnidirectional viewer. In *Technical Report, Bell Laboratories*, Holmdel, NJ, USA, February 1996.
- [53] S. Nayar and V. Peri. Folded catadioptric cameras. In *in Proc. of IEEE International Conference on Computer Vision*, Fort Collins, USA, 1999.
- [54] M. Ollis, H. Herman, and S. Singh. Analysis and design of panoramic stereo vision using equi-angular pixel cameras. In *Technical Report, The Robotics Institute, Carnegie Mellon University*, 5000 Forbes Avenue Pittsburgh, PA 15213, 1999.
- [55] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, 1965.
- [56] A. Paulino and H. Araujo. Multiple robots in geometric formation: Control structure and sensing. In *in Proc. of the 8th Int. Symposium on Intelligent Robotic Systems*, Reading, UK, July 2000.
- [57] S. Peleg, M. Ben-Ezra, and Y. Pritch. Omnistereo: Panoramic stereo imaging. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 23(3):279–290, March 2001.
- [58] W. Press, S. Teukolsky, W. Vetterling, and B. Flannery. *Numerical Recipes in C++, Second Edition*. Cambridge University Press, 2002.
- [59] J. G. Semple and G. T. Kneebone. *Algebraic Projective Geometry*. Clarendon Press, 1998.
- [60] P. Sharkey, D. Murray, S. Vandevelde, I. Reid, and P. Mclauchlan. A modular head/eye platform for real time reactive vision. *Mechatronics*, 3(4):517–535, 1993.
- [61] H. Shum and R. Szeliski. Systems and experiment paper: Construction of panoramic image mosaics with global and local alignment. *International Journal of Computer Vision*, 36(3):101 – 130, 2000.
- [62] Barry Spain. *Analytical Conics*. Pergamon Press, 1957.
- [63] Barry Spain. *Analytical Quadrics*. Pergamon Press, 1960.

BIBLIOGRAPHY

- [64] J. Spletzer, A. Das, R. Fierro, C. Taylor, V. Kumar, and J. Ostrowski. Cooperative localization and control of multi-robot manipulation. In *in Proc. of the IEEE Int. Conf on Intelligent Robots and Systems*, Maui, Hawaii, USA, 2001.
- [65] Jorge Stolfi. *Oriented Projective Geometry*. Academic Press, 1991.
- [66] Peter Sturm. A method for 3d reconstruction of piecewise planar objects from single panoramic images. In *in Proc. of IEEE Workshop on Omnidirectional Vision*, pages 119 – 126, Hilton Head Island, USA, June 2000.
- [67] Peter Sturm. Mixing catadioptric and perspective cameras. In *in Proc. of IEEE Workshop on Omnidirectional Vision*, pages 37 – 44, Copenhagen, Denmark, July 2002.
- [68] T. Svoboda, T. Padjla, and V. Hlavac. Epipolar geometry for panoramic cameras. In *in Proc. of European Conference on Computer Vision*, pages 218 – 332, Freiburg, Germany, 1998.
- [69] G. Taubin. Estimation of planar curves, surfaces and non planar space curves defined by implicit equations, with applications to edge and range image segmentation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 13(11), November 1991.
- [70] Camillo J. Taylor. Video plus: A method for capturing the structure and appearance of immersive environments. *to appear on IEEE Trans. on Visualization and Computer Graphics*.
- [71] R. Vidal, O. Shakernia, and S. Sastry. Distributed formation control with omnidirectional vision based motion segmentation and visual servoing. *IEEE Robotics and Automation Magazine*, to appear 2003.
- [72] W. Wilson, C. Hulls, and G. Belles. Relative end effector control using cartesian position-based visual servoing. *IEEE Trans. on Robotics and Automation*, 12(5):684–696, October 1996.
- [73] N. Winters, J. Gaspar, G. Lacey, and J. Santos-Victor. Omnidirectional vision for robot navigation. In *IEEE Workshop on Omnidirectional Vision*, June 2000.

BIBLIOGRAPHY

- [74] Y. Yagi, S. Kawato, and S. Tsuji. Real-time omnidirectional vision sensor(copis) for vision guided navigation. *IEEE Trans. on Robotics and Automation*, 10, 1994.
- [75] K. Yamazawa, Y. Yagi, and M. Yachida. Obstacle avoidance with omnidirectional image sensor hyperomni vision. In *in Proc. of the IEEE International Conference on Robotics and Automation*, pages 1062 – 1067, May 1995.
- [76] Z. Zhang. Parameter estimation techniques: A tutorial with application to conic fitting. In *INRIA Raport de Recherche n 2676*, October 1995.