

DEPARTAMENTO DE ENGENHARIA ELECTROTÉCNICA E
COMPUTADORES
FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE DE COIMBRA

**NONCENTRAL CATADIOPTIC SYSTEMS
WITH QUADRIC MIRRORS
- GEOMETRY AND CALIBRATION -**



NUNO MIGUEL MENDONÇA DA SILVA GONÇALVES

COIMBRA - FEBRUARY 2008

DISSERTAÇÃO SUBMETIDA À
FACULDADE DE CIÊNCIAS E TECNOLOGIA DA UNIVERSIDADE DE
COIMBRA
PARA A SATISFAÇÃO PARCIAL DOS REQUISITOS CONDUCENTES AO
GRAU DE DOUTOR

TESE EFECTUADA SOB A ORIENTAÇÃO DO
PROFESSOR DOUTOR HELDER DE JESUS ARAÚJO
PROFESSOR CATEDRÁTICO DEEC/FCTUC

Agradecimentos

O trabalho desta dissertação de doutoramento tem o contributo directo e indirecto de muitas pessoas, às quais é mister agradecer.

Em primeiro lugar, quero agradecer ao meu orientador, o Professor Helder, o apoio dado, tanto ao nível científico como ao nível pessoal, na motivação transmitida e no empenho em desenvolver este trabalho. O apoio na procura da precisão da linguagem foi também de grande ajuda. Quero igualmente agradecer a todos os colegas do grupo de Visão, em especial ao João Barreto, pelo interesse demonstrado e pelas conversas de divagação científica, muitas vezes tão úteis na definição e redefinição do rumo do trabalho.

Um agradecimento também ao Dr. Chu-Song Chen e ao Dr. Wen-Yan Chang que prontamente disponibilizaram o código do método por si desenvolvido para a solução do problema da *pose*, tornando possível uma rápida comparação com o meu método de estimação de *pose*.

Ao Instituto de Sistemas e Robótica, ao Departamento de Engenharia Electrotécnica e Computadores da FCTUC e à Fundação para a Ciência e a Tecnologia pelo apoio institucional indispensável à realização do doutoramento.

Quero agradecer a todos os meus amigos pelo apoio e pela amizade demonstrados diariamente, nesta fase tão importante da minha vida. Com medo de me esquecer de alguém importante, agradeço ao Carlos Alberto e ao Carlos David, ao Ricardo (que também me ajudou a fazer uma capa bonita

para a tese), à Gabi, à Tília e ao Ricardo, ao Mário, ao David, à Guida e à Fátima. Não posso, ainda, esquecer os grupos de amigos do Banco de Tempo, com uma nota especial para a Maria Eugénia que me ajudou a rever o inglês de parte importante da tese, e do Choral Poliphónico de Coimbra, dois dos projectos aos quais me dediquei com carinho e dos quais recebi grande retorno em amizade e afecto.

Ao Jorge, por tudo. Por todo o apoio nas componentes essenciais da vida, que aqui não caberiam.

À minha família, numerosa, excelentes pais, grandes irmãos, cunhadas, sobrinhos e sobrinhas, e Taia, pelo apoio constante e pelo afecto superlativo.

Contents

Resumo	xix
Abstract	xxi
I PRESENTATION	1
1 Introduction	3
1.1 Our work	7
1.2 List of main contributions	10
1.3 Outline of the thesis	11
2 The State of the Art	13
2.1 Calibration and geometry of catadioptric systems and non-classical cameras	13
2.1.1 Central catadioptric systems	14
2.1.2 Noncentral catadioptric systems	16
2.1.3 General cameras	18
2.1.4 Other non-classical cameras	20
2.2 Pose of catadioptric systems and the NPnP problem	22
2.2.1 Perspective cameras	23
2.2.2 Catadioptric and other non-classical cameras	25

3	Math tools	27
3.1	Geometric entities and homogeneous coordinates	27
3.1.1	Points	28
3.1.2	Planes	28
3.1.3	Lines	29
3.1.4	Quadrics	33
3.2	Some Geometric Relations in 3D Space	35
3.3	Euler Angles and Rotation Matrices	37
3.4	Quaternions	39
3.5	Intersection of Two Quadrics	41
II	MODELS AND METHODS	47
4	Geometry of Catadioptric Image Formation	49
4.1	Image Formation and the Projection Model	49
4.1.1	Snell's Law	51
4.1.2	Fermat Principle	53
4.2	A New Projection Model	54
4.2.1	Restrictions imposed on the reflection point	55
4.2.2	Computing the reflection point \mathbf{R}	57
4.3	Inverting the projection model	58
5	Method 1: Quadric Mirror Shape Recovery and Calibration of Catadioptric Systems	63
5.1	Problem statement	64
5.2	Apparent Contour	65
5.3	Nonlinear calibration of the catadioptric system	67
5.4	Quadric Pose Estimation	68
5.5	Experiments	73

5.6	Summary and Conclusions	87
6	Method 2: Estimating Parameters of Noncentral Catadioptric Systems using Bundle Adjustment	89
6.1	Problem statement	91
6.2	Projection model	94
6.3	Ray Calibration	98
6.4	Bundle adjustment	99
6.4.1	Initial Estimate	103
6.4.2	Apparent Contour	110
6.5	Experiments	110
6.5.1	Experiments with Simulated Data	110
6.5.2	Real Image Experiments	123
6.6	Summary and Conclusions	128
	Appendix A: Jacobian of Projection Model	130
	Appendix B: Apparent Contour Term	148
7	Method 3: Linear Solution for the Pose Estimation of Non-central Catadioptric Systems	153
7.1	Problem Statement	154
7.2	Pose Estimation	155
7.3	Experiments	160
7.4	Summary and Conclusions	166
III	DISCUSSION	169
8	Discussion and Conclusions	171

List of Figures

4.1	Specular reflection	52
4.2	The light rays reflection and imaging in a catadioptric vision system.	55
5.1	Coordinate systems of the camera, world and mirror and their relative positions.	65
5.2	Real images used.	77
5.3	Reprojection of the 3D points in the image. The left column presents results obtained using the apparent contour and the right column the results obtained without using the contour. In (a)-(b) the initial estimate is the ground truth vector with Gaussian noise added, with standard deviation of 50% of the actual value, in (c)-(d) the initial estimate is a random vector between 0 and 1 (0 \rightarrow 100% of the actual value) and in (e)-(f) the initial estimate is a random vector between -1 and 1 (-100% \rightarrow 100% of the actual value).	85
6.1	World reference frame and the catadioptric system.	93

6.2	Projection model mapping lines in space (given by couples of points A_i and B_i) to pixels in the image. Notice that Snell's law is relaxed so that the incident and reflection angles are not constrained to be equal.	95
6.3	Algorithm for the initial estimate - Mirror/optical axis.	104
6.4	Algorithm for the initial estimate - Center of the hyperbolic mirror.	106
6.5	Algorithm for the initial estimate - Center of the spherical mirror.	106
6.6	Angle error (RMS value in degrees) between the computed and actual incident rays in 3D space. The angle error is computed for an increasing value of the RMS distance between the 4 image pixels used in the computation of the homography.	112
6.7	Results with simulated data . Relative error, or angular error for Euler angles, of the estimated state vector elements, for increasing noise energy added to the initial estimate. The results are plotted in percentage - % and degrees for angles. The parameters whose ground truth value is zero are omitted since no relative error can be computed - ν , θ_2 and B_s . The standard deviation of the noise added is expressed in percentage of the ground truth value.	113
6.8	Cont. - Results with simulated data . Relative error, or angular error for Euler angles, of the estimated state vector elements, for increasing noise energy added to the initial estimate. The results are plotted in percentage - % and degrees for angles. The parameters whose ground truth value is zero are omitted since no relative error can be computed - ν , θ_2 and B_s . The standard deviation of the noise added is expressed in percentage of the ground truth value.	114

- 6.9 Results with **simulated data**. Image error and angle deviation from Snell's Law obtained with the estimated state vectors. These error measures were computed in points not used in the bundle adjustment process. The standard deviation of the noise added is expressed in percentage of the ground truth value. 115
- 6.10 Results with **simulated data**. Relative error, or angular error for Euler angles, of the estimated state vector elements, for increasing noise energy added to the initial estimate and to the image pixel coordinates. The method is tested as a whole. The results are plotted in percentage - % for non angular parameters and degrees for angular ones. The parameters whose ground truth value is zero are omitted since no relative error can be computed - ν , θ_2 and B_s . The standard deviation of the noise added is expressed in percentage of the ground truth value. 118
- 6.11 **Cont.** - Results with **simulated data**. Relative error, or angular error for Euler angles, of the estimated state vector elements, for increasing noise energy added to the initial estimate and to the image pixel coordinates. The method is tested as a whole. The results are plotted in percentage - % for non angular parameters and degrees for angular ones. The parameters whose ground truth value is zero are omitted since no relative error can be computed - ν , θ_2 and B_s . The standard deviation of the noise added is expressed in percentage of the ground truth value. 119

6.12 Results with **simulated data**. Relative error, or angular error for Euler angles, of the estimated state vector elements, for increasing noise energy added to the image pixel coordinates. The initial estimate is obtained by using our automatic algorithm. The method is tested as a whole. The results are plotted in percentage - % for non angular parameters and degrees for angular ones. The parameters whose ground truth value is zero are omitted since no relative error can be computed - ν , θ_2 and B_s . The standard deviation of the noise added is expressed in percentage of the ground truth value. 120

6.13 **Cont.** - Results with **simulated data**. Relative error, or angular error for Euler angles, of the estimated state vector elements, for increasing noise energy added to the image pixel coordinates. The initial estimate is obtained by using our automatic algorithm. The method is tested as a whole. The results are plotted in percentage - % for non angular parameters and degrees for angular ones. The parameters whose ground truth value is zero are omitted since no relative error can be computed - ν , θ_2 and B_s . The standard deviation of the noise added is expressed in percentage of the ground truth value. 121

6.14	Results with simulated data . Image error and angle deviation from Snell's Law obtained with the method as a whole, introducing error in the image pixel coordinates and in the incident rays, correspondent to the tests using our automatic algorithm for the initial estimate that are presented in the set of figures 6.12 and 6.13. These error measures were computed in points not used in the bundle adjustment process. The standard deviation of the noise added is expressed in percentage of the ground truth value.	122
6.15	Real images taken with our experimental setup.	127
7.1	The light rays reflection and imaging in a catadioptric vision system.	155
7.2	Pose transformation between camera and world coordinate systems - H . The world coordinate system can be positioned in the scene object reference frame without loss of generality.	156
7.3	Error analysis tests with simulated data using a pinhole camera and a hyperbolic mirror . Noise was added separately to the input data of the algorithm: image points and 3D data points. Graphics (a) and (b) plot the error for respectively the Euler angles and translation components when noise is added to the structure points and graphics (c) and (d) plot the same error measures for the case where the noise is added to the image point coordinates.	163
7.4	Real images used to estimate the pose of the camera in the world reference frame.	164
7.5	Definition of the terms in which the optimization is made for the nonlinear calibration - the distance d between a 3D point and the back projected incident ray.	165

List of Tables

3.1	Euclidean type of a quadric \mathbf{Q} in terms of its inertia.	42
3.2	Optimal parameterization of the projective quadrics \mathbf{R}_c of inertia different from $(3, 1)$. The parameterized point \mathbf{X}_c satisfy the quadric equation in its canonical form: $\mathbf{X}_c^T \mathbf{R}_c \mathbf{X}_c = 0$. See [32] for the details.	44
5.1	Tests with simulated data for a perspective camera with an off-axis spherical mirror configuration. This table presents median values of the results. Notice that the parameters of the error random distribution (top row) are in relation to the ground truth values of the parameters. GT means ground truth, $N(\cdot)$ means normal error distribution and $\text{Unif}(\cdot)$ means uniform error distribution.	75
5.2	Tests with simulated data. The configuration is made up of an orthographic camera with an aligned parabolic mirror. This table presents median values of the results. Notice that the parameters of the error random distribution (top row) are in relation to the ground truth values of the parameters. GT means ground truth, $N(\cdot)$ means normal error distribution and $\text{Unif}(\cdot)$ means uniform error distribution.	76

- 5.3 Tests with real images obtained from a perspective camera with a **spherical** mirror. Median values of estimated parameters for the quadric mirror in the camera coordinate system - \mathbf{Q}_{cam} and for the pose of the camera in relation to the world reference frame - \mathbf{H} , using the **apparent contour** to reduce the uncertainty. Notice that the parameters of the error random distribution (top row) are in relation to the ground truth values of the parameters. GT means ground truth, $N(\cdot)$ means normal error distribution and $\text{Unif}(\cdot)$ means uniform error distribution. 79
- 5.4 Tests with real images obtained from a perspective camera with a **spherical** mirror. Median values of the estimated parameters for the quadric mirror in the camera coordinate system - \mathbf{Q}_{cam} and for the pose of the camera in relation to the world reference frame - \mathbf{H} , not using the **apparent contour**. Notice that when the apparent contour is not used the algorithm only converges to the solution with good initial parameter estimates. Also notice that the parameters of the error random distribution (top row) are in relation to the ground truth values of the parameters. GT means ground truth, $N(\cdot)$ means normal error distribution and $\text{Unif}(\cdot)$ means uniform error distribution. 80

5.5	Tests with real images obtained from a perspective camera with a hyperbolic mirror. Median values of the estimated parameters for the quadric mirror in the camera coordinate system - \mathbf{Q}_{cam} and for the pose of the camera in relation to the world reference frame - \mathbf{H} , without using the apparent contour . Notice that the parameters of the error random distribution (left column) are in relation to the ground truth values of the parameters. GT means ground truth, $N(\cdot)$ means normal errors distribution and $\text{Unif}(\cdot)$ means uniform error distribution.	81
5.6	Experimental results with real images obtained by a perspective camera with a spherical mirror. Estimated values for the pose of the mirror in relation to the camera - \mathbf{t} and for its radius, using the apparent contour	82
5.7	Experimental results with real images obtained by a perspective camera with a spherical mirror. Estimated values for the pose of the mirror in relation to the camera - \mathbf{t} and for the radius, without using the apparent contour . Notice that when the apparent contour is not used the algorithm only converges to the solution if started with good initial estimates.	83
5.8	Experimental results with real images obtained by a perspective camera with a hyperbolic mirror. Estimated values for the pose of the mirror in relation to the camera - \mathbf{t} and for the radius, without using the apparent contour . Notice that when the Gaussian error introduced is very high, the algorithm does not converge to an useful solution.	83
5.9	Reprojection error of 3D points (in pixels) and the corresponding amplitude error (in percentage) and angle errors (in degrees) of the estimated state vectors.	86

6.1	Standard deviations used to add noise to the image coordinates (in pixels) and to the state vector parameters (in percentage of the ground truth value) in the experiments of the whole method with simulated data.	116
6.2	Final calibration and intermediate initial estimates relative errors in percentage for the hyperbolic configuration. Absolute values are presented for the zero-valued truth parameters. . .	124
6.3	Final calibration and intermediate initial estimates relative errors in percentage for the spherical configuration. Absolute values are presented for the zero-valued truth parameters. . .	125
7.1	Experimental tests using real images acquired by a pinhole camera attached to a spherical mirror . The pose transformation is estimated and the Euler rotation angles and translation elements are listed for the presented and Chen and Chang algorithms. The results are also compared before and after the nonlinear refinement performed to enhance accuracy. Angles are expressed in radians and displacements in <i>mm</i>	166
7.2	Experimental tests using real images acquired by a pinhole camera attached to a hyperbolic mirror . The pose transformation is estimated and the Euler rotation angles and translation elements are listed for the presented and Chen and Chang algorithms. The results are also compared before and after the nonlinear refinement made to enhance accuracy. Angles are expressed in radians and displacements in <i>mm</i>	167
7.3	Comparison between known motion and displacement given by the two poses estimated by the method described. We used the catadioptric system with the hyperbolic mirror. Angles are expressed in radians and displacements in <i>mm</i>	167

Resumo

Nesta dissertação de doutoramento estudamos e analisamos a geometria dos sistemas catadióptricos não-centrais compostos por uma câmara *pinhole* ou ortográfica e um espelho curvo, cuja forma é uma quádrlica não degenerada, incluindo elipsóides, que podem ser esferas, hiperbolóides e parabolóides. A geometria destes sistemas de visão é parameterizada, analisando o fenómeno de formação da imagem, e é composta pelos parâmetros intrínsecos da câmara, os parâmetros da superfície do espelho e a posição e orientação da câmara em relação ao espelho e ao sistema de referência do mundo. A formação da imagem é estudada numa perspectiva puramente geométrica, focando principalmente o modelo de projecção e a calibração do sistema de visão.

As principais contribuições deste trabalho incluem a demonstração de que num sistema catadióptrico não-central com um câmara em perspectiva e uma quádrlica não degenerada, o ponto de reflexão na superfície do espelho (projectando na imagem qualquer ponto 3D do mundo) pertence a uma curva quártica que é dada pela intersecção de duas superfícies quádrlicas. O correspondente modelo de projecção é também desenvolvido e é expresso através de uma equação não linear implícita, dependente de um único parâmetro.

Relativamente à calibração destes sistemas de visão, foi desenvolvido um método de calibração, assumindo o conhecimento dos parâmetros intrínsecos da câmara em perspectiva e de um conjunto de pontos 3D expressos em coordenadas locais (estrutura 3D do mundo). Informação acerca do contorno

aparente do espelho é também usada para melhorar a precisão da estimação. Um outro método de calibração é proposto, assumindo uma calibração prévia do sistema no sentido de um modelo geral de câmara (correspondências entre pontos na imagem e raios incidentes no espaço).

Adicionalmente, a posição e orientação (*pose*) da câmara em relação ao espelho e ao sistema de referência do mundo são estimadas usando métricas algébricas e equações lineares (escritas para um método de calibração que também é apresentado). Considera-se a câmara como pré-calibrada.

São desenvolvidas e apresentadas experiências com simulações extensivas e também com imagens reais de forma a testar a robustez e precisão dos métodos apresentados.

As principais conclusões apontam para o facto de estes sistemas de visão serem altamente não lineares e a sua calibração ser possível com boa precisão, embora difícil de alcançar com precisão muito elevada, especialmente se o sistema de visão tem como objectivo aplicações direccionadas para a precisão. Apesar disso, pode observar-se que a informação da estrutura do mundo pode ser complementada com informação adicional, tal como o contorno aparente da quádrlica, de forma a melhorar a qualidade dos resultados de calibração. Na verdade, o uso do contorno aparente do espelho pode, por si, melhorar drasticamente a precisão da estimação.

Abstract

In this PhD thesis we study and analyze the geometry of noncentral catadioptric systems composed by a pinhole or orthographic camera and a non-ruled quadric shaped mirror, that is to say an ellipsoid, which can be a sphere, a hyperboloid or a paraboloid surface. The geometry of these vision systems is parameterized by analyzing the image formation and is composed by the intrinsic parameters of the camera, the parameters of the mirror surface and the poses of the camera in relation to the mirror and to the world reference frames. Image formation is studied in a purely geometrical way, focusing mainly on the projection model and on the calibration of the vision system.

The main contributions include the proof that in a noncentral catadioptric system with a perspective camera and a non degenerate quadric the reflection point on the surface (projecting any given 3D world point to the image) is on the quartic curve that is the intersection of two quadrics. The projection model related to the previous definition of the reflection point is also derived and is expressed as an implicit non linear function on a single unknown.

In what concerns the calibration of these vision systems, we developed a calibration method assuming the knowledge of the intrinsic parameters of the perspective camera and of some 3D points in a local reference frame (structure) . Information about the apparent contour is also used to enhance the accuracy of the estimation. Another calibration method is proposed,

assuming a previous calibration of the system in the sense of a general camera model (correspondences between image points and incident lines in space).

Additionally, the camera-mirror and camera-world poses are estimated using algebraic metrics and linear equations (derived for a calibration method that is also presented). The camera is considered to be pre-calibrated.

Experiments with extensive simulations and also using real images are performed to test the robustness and accuracy of the methods presented.

The main conclusions are that these vision systems are highly non linear and that their calibration is possible with good accuracy but difficult to achieve with very high accuracy, specially if the vision system is aimed at being used for accuracy-driven applications. Nevertheless it is observed that structure of the world can be complemented with some additional information as the quadric apparent contour in order to improve the quality of the calibration results. Actually, the use of the apparent contour can dramatically improve the accuracy of the estimation.

Part I

PRESENTATION

Chapter 1

Introduction

Cameras are system devices that acquire images of the world. Often images are bi-dimensional representations of tri-dimensional scenes where the transformation of the visual information involved always implies loss of information. Mathematically this transformation is expressed as a projection from a higher dimensional space into a smaller one.

The pinhole camera model widely used in computer vision, which expresses the visual transformation as a perspective projection, is still the most important camera model. This simple and linear model has some limitations, however.

To overcome the limitations of the vision systems made up of perspective cameras, several different solutions involving new design of vision systems have, in the recent years, been proposed and studied, usually aiming to attain some specific task or optimality in some particular characteristic. Even if the usual pinhole camera is perfectly suitable for a wide variety of applications, in several others it can be advantageously replaced by more complex optical setups involving the combination of mirrors and lenses.

Wider fields of view (since smaller ones are one of the most expressive geometrical limitations of conventional cameras), for example, are extremely

useful for applications such as surveillance and tracking, among many other.

It has been known since ancient times that mirror (initially made of metal) and glass have the property of changing and redirecting light rays, by reflecting or refracting their energy. Ancient civilizations were already aware of the focal properties of some specular surfaces and used them. Today we suppose that the discovery of glass was incidentally made about 5000 B.C. by the Phoenicians (described by the Roman historian Pliny) and that its focal properties were used by ancient Greeks to induce fire. The first known study on the human visual system was made by Ibn al-Haitham in his Book of Optiks, where he describes how the lens in the human eye forms an image on the retina [60]. Later technology allowed mankind to use mirrors and lenses in precisely controllable deviations of the light.

This also allowed the construction of vision systems that combining cameras, mirrors and lenses, could achieve wider fields of view. For that purpose the directions of the light rays have to be changed in such a way as to guarantee that most of the scene can be imaged by a single sensor. The geometric properties of such vision systems depend on whether lenses or mirrors (or combinations of both) are used. The development of those new configurations for vision sensors also implied that several problems had to be tackled, mainly those concerning the development of methods and models for projection, calibration, 3D reconstruction, motion and others. Systems that use mirrors and cameras are called catadioptric and they have recently been exhaustively studied.

Catadioptric vision systems can be divided into two types depending on whether the projection is central or not, in other words, depending on whether all incident light rays intersect each other in a unique viewpoint or not. The locus of viewpoints depend only on the geometric properties of the camera, mirrors, lenses and their relative positioning.

Specifically, it has been shown by Baker and Nayar [6] that for quadric

mirror catadioptric systems, the central projection can be obtained only for a particular position of the camera optical center, usually the focus of the quadric. In those cases there exists a closed form expression for the projection model [12, 47]. However, for the general case and when this constraint is relaxed, the projection is non central which implies that the light rays do not intersect at an effective single viewpoint.

Noncentral vision systems do not generally have a projection model. As a result closed form expressions relating 3D world point coordinates to their corresponding image coordinates do not exist whilst they are known for the case of central systems. The projection through a noncentral catadioptric systems has hence been solved by using either the Snell's Law or the Fermat Principle [23], that provide implicit multivariate expressions.

The existence of an explicit projection model is very important for vision systems since the ability of relating not only a pixel with a light ray direction (also possible in noncentral cameras, provided that the system parameters are known) but also a 3D world point to a specific pixel allows a much more complete theory. In addition, such a model also allows the use of several algorithms based on those geometric relationships. In this context, the calibration is easier to perform in central systems (regardless of their type and geometry) than in noncentral vision systems.

In fact this is the reason why noncentral vision systems are used mainly in applications that do not require extreme accuracy, such as navigation, tracking and visualization. Central vision systems, whether they are perspective, catadioptric or other, are usually also used in accuracy-driven applications, such as 3D reconstruction, motion, structure from motion, or distance and angle measurement.

To overcome the non existence of a projection model for some types of vision devices, a new model of cameras has been proposed, namely the generalized cameras [55, 124]. This class of cameras, an abstraction of the

image formation phenomenon, are also called black-box cameras and they associate each pixel with a direction in space. Hence, practically all cameras (central or not) can be described by a general model. Calibration of those vision systems results in a list of correspondences between each pixel in image and a line in space.

Another important problem in computer vision besides geometry and design of catadioptric systems is the estimation of the camera pose (composed by location and orientation) in relation to the world, also called extrinsic parameters. This estimation is a definitely important problem both in computer vision and in robotics being relevant for several applications, particularly motion estimation, structure from motion, robot navigation, self-localization, object recognition, head and body posture and many others.

Several algorithms exist to estimate the pose of a camera in the world, for central and noncentral projection models, with different types of assumptions and using a single or multiple images.

The classical approach to the problem of estimating the position and orientation of the camera in relation to the world frame is the perspective n -point (PnP) problem. The problem was originally formulated by Fischler and Bolles [37] as the calculation of the distance and orientation of some points to the camera optical center (their coordinates in camera frame). Another understanding of the problem is formulating it as the estimation of the screw transformation matrix between the camera and world coordinate systems. The world coordinate system is sometimes called local or object coordinate system without loss of generality. Those two formulations of the problem of pose (as a PnP problem or as a screw transformation matrix) are closely related to each other and many times are indistinguishable in literature [69].

Several solutions for both have been presented until now. Initially the approach tends to focus on geometrical properties of triangles or quadruples

in space and of their projection through the optical system into image. Planar and non planar solutions have been formalized. Some different approaches either using multiple views or appearance models were then proposed to substitute or complement the initial analytical derivations. As for to the features used, the most popular ones are points. However, lines, planes, circles, ellipses and curved shapes have also been used in the pose estimation, providing a vast number of estimation methods.

For non perspective cameras, there are recent works either in general cameras or in more specific noncentral catadioptric systems. The non existence of an explicit projection model is again limiting. Chen and Chang [21] presented the solution of the non-perspective 3-point problem (NPnP) for a non central general camera assuming the knowledge of a direction in space corresponding to each pixel (calibrated camera in the sense of a general model). Using a different algebraic approach, Nister [100] and Stewénus et al. [120] also solved the problem of the pose estimation for generalized cameras considering a pre-calibrated general camera and providing algebraic constraints to the problem. There are some other geometric solutions but, due to the nature of the projection, they are all non linear and usually computed by iterative optimization.

1.1 Our work

In this thesis we are interested in catadioptric systems composed by a conventional perspective camera and curved mirrors expressed by non degenerate quadric shapes. These mirrors expressed by full rank quadrics are very popular in omnidirectional vision and include surfaces with conical section as hyperboloids, paraboloids and ellipsoids (being the sphere a particular case of the last one).

We are particularly interested in all noncentral configurations of the vi-

sion system so that the camera positioning in relation to the specular surface will be unconstrained.

Our work presented here can be divided into four different parts.

The first one is about the geometry of image formation in noncentral catadioptric systems. The projection model understood as the mapping between a 3D world point in space and the corresponding image pixel is specifically studied. The reflection of a light ray through the specular surface and its projection into the image is explained by two different principles of optics: the Snell's Law and the Fermat Principle. None of them, however, provide closed form solutions to this projection in an unconstrained noncentral catadioptric vision systems. We studied this topic and we could not find the wanted closed form projection model. However, some insight is given into the problem and a novel alternative framework for the geometry of image formation is proposed. This method to project a 3D point into image expresses the corresponding pixel as the solution of a nonlinear function in a single parameter, whereas the use of either the Snell's Law or the Fermat Principle leads to nonlinear functions that depend on several unknowns.

The second and third parts of the work concern the problem of calibrating the vision system. In the second one we use in all derivations a fully parameterized system and consider that the pinhole intrinsic parameters are known. By applying nonlinear optimization methods to the problem, our framework is able to estimate the quadric shaped mirror parameters in camera coordinates and the transformation matrix between the camera and world coordinate systems (camera-world pose). In a subsequent step, the mirror is expressed in its canonical form (allowing for its classification) and the transformation matrix between the camera and mirror reference frames (camera-mirror pose) is estimated. We prove that the use of the apparent contour of the quadric mirror dramatically improves the accuracy of the estimations.

The other calibration method we propose is a new method to calibrate the intrinsic and extrinsic parameters of a generalized catadioptric system (considered to be noncentral, although it is also suitable to the central projection case). This two-steps method firstly calibrates light rays by providing correspondences between incident lines in space and pixels in an arbitrary world reference frame (in the sense of a general camera model calibration). The second step hence calibrates the intrinsic parameters of the pinhole camera, the coefficients of the mirror expressed by a quadric, the position of the optical center of the camera in the world reference frame and its relative orientation (as in the method previously explained, the system is thus fully parameterized). A projection model relaxing Snell's Law is derived and the deviations from the Snell's Law and the image reprojection errors are minimized by means of bundle adjustment using explicitly the derived jacobian. Information about the apparent contour of the mirror can also be used to reduce the uncertainty in the estimation by either reducing the dimension of the problem or by introducing a new term in the cost function.

Finally, the fourth part of the work is related to the pose estimation. The approach presented is fully linear and assumes that the intrinsic parameters of the system are available (possibly calibrated using the previous methods). It is based on the derivation of a linear constraint in the pose transformation elements (nine elements of the rotation matrix and three elements of the translation vector). This analytical constraint is based on the projection model proposed in the first part of the work, that is to say, that the reflection point (on the mirror surface) belongs also to an analytical quadric whose coordinates are dependent on the mirror, the 3D world point to project and the optical center. The orthogonality of the rotation matrix is enforced after this estimation to approximate the final rigid solution. We hence proved that if the camera is considered to be previously calibrated, its pose in the world reference frame can be estimated using an algebraic linear method.

We emphasize that the three methods presented complement each other in such a way that it is possible to fully calibrate a noncentral catadioptric system with quadric mirrors and calculate the pose of the vision system in relation to the world with accuracy. The system is first calibrated by the first method using only the pinhole camera parameters and the result of this calibration gives the input data needed for the second method to improve the accuracy of the full system calibration. These results can therefore be used to estimate with high accuracy the pose of the camera in relation to the world.

1.2 List of main contributions

The main contributions to the computer vision area of our work presented throughout this thesis are listed below.

1. The proof that in a noncentral catadioptric system with a perspective camera and a non degenerate quadric the reflection point on the surface that projects any given 3D world point to image is in the quartic curve that is the intersection of two quadrics: the quadric mirror itself and an analytical quadric that depends on the mirror, the optical center and the projected 3D world point.
2. The projection model related to the previous definition of the reflection point. This projection model can be expressed as an implicit nonlinear function on a single unknown.
3. A calibration method for the same noncentral catadioptric vision systems assuming only the intrinsic parameters of the perspective camera and some 3D points in a local reference frame.
4. A calibration method for the same noncentral catadioptric vision systems assuming a previous calibration of the system in the sense of a

general camera model.

5. A camera-world pose estimation method for the same noncentral catadioptric vision systems using an algebraic metric and linear equations. The camera is considered to be calibrated.

1.3 Outline of the thesis

We start by presenting in the next chapter the state of the art in catadioptric vision systems. It is not our aim to be exhaustive in the description of this scientific topic but rather present the main problems and contributions made to it in recent years. The author is aware of the fact that every summary is always biased by the understanding of its writer.

In chapter 3 we present some mathematical tools essential to the understanding of the thesis work, mainly in geometry and analytical algebra. This completes the first part of the document.

In the second part, we present in every chapter a different contribution. In chapter 4 the geometry of catadioptric image formation is analyzed. The novel projection model is derived and proposed. Chapters 5 and 6 present the two calibration methods proposed and then in chapter 7 we describe the framework for the estimation of the world-camera pose linearly.

Finally, in the third and last part of the thesis, we discuss the main contributions, and draw final conclusions. Extensive bibliography is also provided.

Chapter 2

The State of the Art

The aim of this chapter is to review the state of the art in catadioptric vision systems, focusing mainly two aspects: calibration and pose. We mean by calibration the process of estimating the intrinsic parameters of camera, reflecting surface and their relative positions. By pose we mean the relative position of the camera and world coordinate systems. Due to the nature of the two problems, we opted to present the state of the art of both separately. Some other items are also mentioned and discussed as geometry, stereo, reconstruction, structure from motion for catadioptric and also perspective and non-perspective cameras.

2.1 Calibration and geometry of catadioptric systems and non-classical cameras

Since vision systems can be composed in several different ways as, for instance, catadioptric systems that are composed by a perspective or orthographic camera and a specular mirror or fish-eye systems that are composed by a perspective camera and a fish-eye lens, or even systems composed by

several mirrors, lenses and cameras, there have been published a multitude of frameworks providing calibration of these systems. This calibration can be obviously understood as the establishment of the correspondences between image pixels and scene points, irrespective of the existence or not of an explicit projection model.

The use of catadioptric and dioptric systems in computer vision arose in the late 70's although it was only in the decade of 90 that it became popular. The advantages of wider fields of view was first appreciated by the roboticists that built many panoramic vision systems based on fish-eye lenses, conic mirrors or even based on perspective swiveling cameras.

As explained in the introduction, the fact that the projection is central or noncentral plays a key role in the calibration and accuracy-driven applications in catadioptric vision. Furthermore, generic camera models that describe the imaging irrespective of the projection model are also very important in computer vision since they provide strong freedom in the designing and use of cameras, mirrors and lenses. Although it is extremely difficult to establish frontiers in knowledge fields, and hence in classification, since they always overlap each other, we opted to divide the state of the art for these different types of cameras.

2.1.1 Central catadioptric systems

Although several catadioptric vision systems existed before (the most representative of them are [64, 97, 137, 138]), the first attempt to provide a geometric framework for the image formation on a single viewpoint catadioptric system is due to Baker and Nayar [6, 7, 98]. They formalized the condition to obtain a central projection with a perspective camera and a reflecting mirror surface (catadioptric system). Nayar and Peri [99] also presented folded catadioptric systems, composed by a perspective camera and two or more mirrors. The conditions to guarantee central projection are then known

and thus it is possible to apply to those systems the theoretical advances in computer vision made so far.

Two other approaches to a unifying theory for central catadioptric vision systems were presented by Geyer and Daniilidis [43, 47] and Barreto and Araújo [11, 12]. The former provides a projective geometry framework for catadioptric systems demonstrating that the image formation of any central catadioptric system can be decomposed into a two-step projection through a sphere. On the other hand, the latter shows that the projection in any central catadioptric system is made up of three functions namely a linear function mapping the world into an oriented projective plane, a nonlinear transformation between two oriented projective planes and a collineation in the plane. This framework was subsequently improved to include also dioptric systems (systems with cameras and lenses with or without radial distortion) as long as they maintain a unique viewpoint [9, 10].

These frameworks quickly allowed the appearance of different calibration theories using geometric features like in [139] (Ying and Hu, 2004) by searching projective invariants in the projection of lines and spheres, in [15] (Barreto and Araújo, 2005) by studying the geometrical properties of the projection of lines through catadioptric systems, in [34] (Fabrizio et al., 2002) by using the inner and outer contour of the mirror and in [89] (López-Franco and Bayro-Corrochano, 2006) by applying conformal geometric algebra to the calibration using not only points, lines and planes but also point pairs, circles and spheres.

In spite of the fact that these studies provide calibration for general central catadioptric systems, the one that has been probably more deeply studied and used so far is the paracatadioptric system which is composed by a orthographic camera (or perspective one with a telecentric lens) and a parabolic shaped mirror. These systems have the advantage of providing central projection irrespective of the vertical location of the camera in rela-

tion to the mirror. Geyer and Daniilidis [48] provided calibration using lines, as well as Barreto and Araújo [13,16]. Structure from motion [44] (Geyer and Daniilidis, 2001) and stereo rectification [46] (Geyer and Daniilidis, 2005) are also formalized for paracatadioptric vision systems.

Multiple view geometry for central catadioptric systems was also formalized in the Epipolar framework by Svoboda and Pajdla [125]. They proved that the epipolar constraint takes the shape of a conic in image and they derive the corresponding conic for all types of central catadioptric systems with conic section mirrors (hyperboloids, paraboloids and ellipsoids). Other contributions for Epipolar geometry by deriving explicit expressions for the Fundamental Matrix are proposed by Geyer and Daniilidis [45] and also by Gupta and Daniilidis [57] by providing calibration from the fundamental matrix in planar motion for parabolic shaped mirrors.

Other recent and important works in the calibration and geometry of central catadioptric vision system include those presented by Daniilidis and Makadia [25] for optical flow computation, Demonceaux and Vasseur [29] who propose a new method using Markov Random Fields for segmentation, image restoration and motion detection, Lin and Bajcsy [86] who present a central catadioptric viewer using a conical mirror, Hicks, Millstone and Daniilidis [62] for the tuning of a two-folded mirror system that performs a given predetermined single viewpoint projection and Scaramuzza, Martinelli and Siegwart [113] that use a Taylor series expansion to approximate the projection model and applies a nonlinear minimization method to calibrate the corresponding central catadioptric system.

2.1.2 Noncentral catadioptric systems

Noncentral catadioptric vision systems are all vision systems composed of a camera and a reflecting mirror whose projection has multiple viewpoints. Due to some advantages over the single viewpoint systems and since the cen-

tral projection is generally obtained for a very particular relative positioning of the camera and the reflector or lens, the study of noncentral projection systems became more and more pertinent to the computer vision community. For a survey study on noncentral cameras before 2000, we suggest the reading of [8] (Bakstein and Pajdla, 2000).

As in central catadioptric systems, before the first attempts to study the geometry and calibration of general noncentral ones, there existed some proposals mainly in the domain of robotics for navigation purposes.

Swaminathan, Grossberg and Nayar [126, 128] presented the geometric study of noncentral catadioptric systems composed of quadric shaped mirrors that include the most used reflector surfaces (hyperboloids, paraboloids and ellipsoids). The locus of all viewpoints is proved to be the caustic surface that univocally completely characterizes the vision system.

Micusik and Pajdla [94] presented a calibration method for the same noncentral catadioptric systems by first approximating the projection to be central and then performing a refinement on the actual noncentral projection. Autocalibration and 3D reconstruction are addressed by using two views of the same vision system. The theory here proposed is extended to include all wide circular field of view cameras in [95] in a two-view framework. The epipolar geometry is studied and used in autocalibration, 3D reconstruction and structure-from-motion estimation.

Bundle adjustment was also proved to be suited to calibration of noncentral catadioptric systems. Lhuillier [84, 85] presented a framework for the calibration of noncentral catadioptric systems using bundle adjustment. Numerical computations of the jacobian are used.

Conic fitting is used by Mashita, Iwai and Yachida [93] to calibrate noncentral catadioptric systems. Since there is some ambiguity in their estimation, they propose a selection method to choose the best set of parameters that better describe the vision system.

Using another approach, Kannala and Brandt [76] present a generic framework that includes not only dioptric and catadioptric wide angle cameras but also conventional ones. They use a projection model composed of two distortion terms: one in the radial direction of the image and the other in the tangential direction.

Lines are also used to calibrate an off-axis catadioptric sensor by Caglioti et al. [19]. They use the apparent contour of the mirror and the parameters estimated are the intrinsic parameters of the pinhole camera, the mirror shape and the camera-mirror pose.

Several applications of noncentral catadioptric geometry have been studied and presented in the last years. Caglioti and Gasparini [20] presented a method to locate 3D lines by a single image.

Differently shaped mirrors are also used in noncentral catadioptric systems. Spacek [119] proposes a conical mirror and an unwarping function to rectify the image. Stereo configurations are also studied to provide range maps for robot navigation. Fiala and Basu [36] present a two-lobe non-conical section mirror that is used as a two view stereo camera which is proved to reconstruct 3D models of polyhedral objects. Hicks [61] also propose new mirror design techniques to perform a predetermined projection. As this problem doesn't have a general solution, the method proposed uses an approximation by means of a numerical solution. The projection model in catadioptric sensors is addressed.

2.1.3 General cameras

General cameras include all vision systems. They are important to catadioptric vision systems since they provide useful theory on calibration and several other applications. The general model of cameras was proposed by Grossberg and Nayar [55]. They define raxels as virtual sensing elements. Using a raxel for each direction in space it is possible to establish the correspondence

between a pixel in the image and an incident light ray direction. The set of all correspondences is called calibration of the system and is achieved using structured light patterns. The raxels proposed include geometric, radiometric and optical properties.

Pless [106] presents this model of general cameras to derive epipolar constraints. Sturm and Ramalingam [124] use a geometric simplification of the general imaging model and propose a method to calibrate these generic camera models using three views. The basic idea is that three points that are imaged in the same pixel are collinear in 3D world reference system. The two screw transformations between the camera in first and second and first and third positions can then be estimated by a high-dimensional lifted linear system. Ramalingam, Sturm and Lodha [111] present some improvements and the specific formalization for central cameras [110].

Structure from motion was studied and presented by Ramalingam, Lodha and Sturm [109]. The approach presented allows to reconstruct scenes from pre-calibrated images possibly taken by cameras of a different type (cross-cameras).

Sturm [122] also presents a multi-view study on general camera models. Several configurations are analyzed and the foundations of multi-view geometry of completely noncentral cameras are derived. The equivalent to fundamental and essential matrices, trifocal and quadrifocal tensors in perspective cameras are derived for this noncentral camera model.

Recently, Dunne, Mallon and Whelan [31] addressed the problem of the calibration of a general camera model with the additional assumption of its centrality. This assumption reduces the complexity of the generic calibration method. Polynomials are also used to estimate the pose of the sensor linearly.

2.1.4 Other non-classical cameras

In recent years a multitude of non-classical cameras have been proposed, usually optimal in a specific characteristic and aimed to achieve a given task. We don't intend to be exhaustive in the citation of non-classical cameras, but rather intend to make a review of the most representative ones, also important to the understanding and study of noncentral catadioptric vision systems.

Panoramic images composed of mosaicing images taken from a swiveling perspective camera have been used in many applications. If perfect alignment is achieved in rotating the camera around its optical center, the system has central projection. Ishiguro, Yamamoto and Tsuji [72] constructed a panoramic vision system using a rotating camera to construct a map of an indoor environment and acquire the range of objects using binocular information. If stereo configurations are required, usually to estimate depth information, the projection is noncentral. Peleg, Erza and Pritch [105] have used two swiveling cameras for panoramic imaging and have studied their geometry and viewpoint locus (caustic surface). Equivalence to compositions of cameras and lenses is also analyzed. Jiang, Sugimoto and Okutomi [74] presented an alternative panoramic configuration for dense depth estimation.

Nalwa [97] presented a pyramidal mirror to acquire a wide field of view and that guarantees a single viewpoint with as many cameras as the faces of the right mirror-pyramid, that is, four cameras in the common configuration. Images from the individual cameras are concatenated to yield a 360 degrees wide panoramic field of view.

Oblique cameras are noncentral cameras whose incident rays don't intersect each other such that a point is imaged only once. Pajdla [103, 104] presented oblique cameras as a generalization of panoramic, pushbrooms, catadioptric systems and other noncentral cameras. Multi-view geometry

is derived for this class of vision systems. Cross-slit (X-slit) cameras are a particularization of oblique cameras where all rays pass through two general skewed lines in space. Their geometry is studied and proposed by Feldman, Pajdla and Weinshall [24]. Their main application is in generating panoramic views from single translating pinhole sequences.

Planar mirrors are also used to construct virtually a stereo configuration to rectify images and estimate depth. Gluckman and Nayar [49] proposed this vision system and studied the underlying geometry.

Swaminathan, Grossberg and Nayar [127] also designed a mirror to achieve a given projection function by minimizing image error. The method presented is linear and is able to find the best solution when the mirror design has no exact physical realization.

Another type of cameras include abstractions of the real image formation and modeling by radial distortion, usually including in this design perspective conventional cameras, dioptric system with fish-eye or other lenses, central and noncentral catadioptric systems and general wide field of view cameras. Barreto and Daniilidis [17] have presented multi-view geometry analysis for these cameras. Tardiff, Sturm and Roy [129] have also presented a self calibration method for these cameras.

Kuthirummal and Nayar [79] proposed the so called radial imaging systems. Single- and multi-view geometry are derived and applied to recover the shape of convex objects by using two images. Texture map and reflectance properties are also addressed.

Rolling-Shutter cameras are those whose image acquisition is not made in a single shot, that is to say, it is not instantaneous. They are cheap cameras and usually hand-held. Geyer and Daniilidis [46] proposed a calibration method for these cameras. They prove that if some conditions are met they can be modeled by a X-slit camera.

Tisse, White and Hicks [131] constructed a multidirectional optical sensor

for navigation. The frontal part of the sensor is equivalent to a conventional perspective camera whereas the non central part of the sensor is composed of four slopes or a pyramid. Geometry of the sensor is studied and single viewpoint is achieved by designing the slope of the pyramid.

Using a dual mirror pyramid Hua, Ahuja and Gao [70] have also used a pyramidal mirror to acquire wide field of view images. Geometry of the system is analyzed and designed to maximize the field of view.

Recently Kuthirummal and Nayar [78] presented the geometrical framework for vision systems composed by flexible mirror. The calibration is performed using the contour of the mirror and by approximating its deformable shape by a product of splines.

Many other non-classical cameras exist since the design of vision systems is highly flexible and new sensors can always be designed to achieve a given objective and serve a specific task.

2.2 Pose of catadioptric systems and the NPnP problem

Pose estimation and the well-known perspective n-point (PnP) problem are not generally the same problem, as shown by Hu and Wu [69]. Despite this fact those two problems are generally indistinguishably defined in literature. The former is the estimation of a transformation matrix - the pose (three rotations and three translations) - between camera and object reference systems whereas the latter, the PnP, is the estimation of the distance of some control points (whose relative positions in the object reference frame are known) to the camera center. However, since both problems are very close to one another we treat them indistinguishably as in literature.

2.2.1 Perspective cameras

We firstly review the pose problem in the conventional perspective cameras.

Fischler and Bolles [37] used cosine laws to determine the distance between the camera and some points in a geometrical approach. This formulation was thereafter linearized by Quand and Lan [107] and later by Ameller et al. [3]. Hung, Yeh and Harwood [71] proved that four coplanar points with known coordinates in an object reference frame are sufficient to determine their coordinates in relation to the camera. Linnainmaa, Harwood and Davis [87] presented another pose estimation algorithm that uses triangle pairs to estimate the screw transformation by first computing candidate position vectors and then the rotation ones. The best solution is chosen by checking geometrical consistency. Horaud et al. [66] presented an analytical solution to the perspective 4-point problem.

Using lines rather than points, Chen [22] proposes the estimation of pose. The solutions are closed-form and are based on a polynomial approach. On the other hand, conics are used by Ma [90] for the estimation of the pose where it is proved that two conics are sufficient to estimate it and if the conics are coplanar, closed-form solutions are derived and presented.

DeMenthon and Davis [27] propose two alternative methods by using only three points and also four non coplanar points [28]. Both methods are based on approximations of the perspective projection. The former addresses the weak-perspective, the paraperspective and the orthoperspective projection whereas the latter uses an orthoperspective projection. Horaud et al. [108] present a framework to approximate the full perspective to the paraperspective projection. All approximations are refined by nonlinear methods to achieve convergence on the full perspective projection. Claiming accuracy for distant objects, Oberkamfot, DeMenthon and Davis [102] presented an iterative method for pose estimation using four or more coplanar points.

Araújo, Carceroni and Brown [5] improved the accuracy of the work previously presented by Lowe [88] to estimate the pose in a fully projective formulation.

Joseph [75] presents a linear solution to the pose estimation in two dimensions to use as a supporting framework to the 3D case. Ji, Costa, Haralick and Shapiro [73] use alternative geometric feature such as points, lines and ellipse-circle pairs for the estimation of pose. Tarel and Cooper [130] also use different features rather than only points. They propose the use of algebraic curves described as polynomials.

Ansar and Daniilidis [4] presented new linear solutions to the PnP problem extending it to include lines. The pose can then be computed by using n points or n lines. The solution is tested in an augmented reality application.

Symmetry is used by Hong et al. [65] for the geometry analysis and derivation. They point out that the symmetry is capable to facilitate the pose dramatically as well as the 3D reconstruction and calibration of a visual system. On a different approach, Davis et al. [26] present a correspondence free algorithm to estimate the pose using nonlinear methods.

In recent years the problem of pose estimation has still been interesting the scientific community. Several works are still being presented and proposed for pose estimation. The most important of them are those presented by Schweighofer and Pinz [116] using a planar target and exploiting the ambiguities of coplanar configurations of four points, Furukawa et al. [40] by using outlines of curved surfaces rather than point correspondences, Nister and Schaffalitzky [101] where four points in two calibrated views are used and critical configurations are addressed, Kazhdan [77] and Rosenhahn et al. [112] that use 3D shapes instead of features to track pose of objects.

Moreno-Noguer et al. [96] presented an accurate non-iterative method for the PnP problem with four or more points. The complexity of the presented method grows linearly with the number of points, which is the main improve-

ment in relation to the existing methods (whose computational complexity is of the fifth order or more). The main idea is to project the points as combinations of four virtual points and hence the estimation is focused on the coordinates of these points in the virtual reference frame.

Some other important works in the pose estimation of perspective cameras include [41, 56, 66–68].

2.2.2 Catadioptric and other non-classical cameras

The pose has been also studied for non perspective cameras and recently the pose in catadioptric vision systems has interested the community.

Aliaga [1, 2] calibrated catadioptric systems in robotic environments for pose estimation. Makadia and Daniilidis [91, 92] have studied the estimation of the 3D rotation of the camera reference system using central catadioptric sensors. They use a generalization of the Fourier Transform and the projection mapping on the sphere for finding invariants of the rotation parameters. Another approach using geometric invariants is due to Ying and Hu [139]. Barreto and Araújo [14] also addressed this problem by providing a selection framework for the camera system in the world reference frame.

The partial estimation of rotation (roll and pitch angles) is proposed by Demonceaux et al. [30] by using the projection of the horizon in the image (line at infinity) in a central catadioptric system.

Uncertainties in pose estimation are dealt with by Gebken, Tolvanen and Sommer [42] for a central catadioptric system. They use line-plane correspondences for geometry and a stochastic method to deal with the uncertainties.

On the other hand, in noncentral projection and general cameras the pose estimation is very recent.

Fabrizio and Devars [35] analyze the non-perspective n -point problem in catadioptric systems in a totally analytical approach. Nonlinear closed-form solution are presented for noncentral projection systems and simplified for

central ones.

With a different analytical approach, Chen and Chang [21] propose an exact solution for three points and an approximate solution with minimal error for more than three points. This method is applicable to generalized noncentral cameras but can also be applied to perspective ones. Nister [100] and Stewénius et al. [120] also solved the problem of the pose estimation for generalized cameras considering a pre-calibrated general camera and providing algebraic constraints to the problem.

Sturm and Bonfort [123] addressed the problem of pose estimation of an object by analyzing reflections on two or three planar mirrors with unknown relative positions using a single image.

Chapter 3

Math tools

There are several mathematical entities, relations and formalisms that are used in engineering. Engineering can actually be regarded as the application of the principles of physics to reality, expressed and described by all mathematical frameworks. Computer vision and particularly panoramic vision is extremely related to the branch of mathematics called geometry. In this chapter, we give some insights into the most used mathematical tools in this work, without the intention to explain every detail, but only the most important to understand the text. It is intended to be a consulting section rather than a textbook for geometry.

3.1 Geometric entities and homogeneous coordinates

Geometric entities as points, planes, lines and quadrics are here represented in homogeneous coordinates as usual in projective geometry, rather than in Cartesian coordinates. The point spaces used are P^2 in the image plane and P^3 in 3D space.

3.1.1 Points

- Points in P^2 are represented algebraically by a 3×1 vector in homogeneous coordinates. They are represented by $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}^T$.
- Points in P^3 are represented algebraically by a 4×1 vector in homogeneous coordinates. They are represented by $\mathbf{X} = \begin{bmatrix} X_1 & X_2 & X_3 & X_4 \end{bmatrix}^T$.
- Any vector $\lambda\mathbf{x}$ in P^2 or $\lambda\mathbf{X}$ in P^3 , where λ is a non-zero real scalar, represents the same point.
- Any transformation in P^3 , represented by matrix \mathbf{H} , is such that $\mathbf{X}' = \mathbf{H}\mathbf{X}$.

3.1.2 Planes

- A point \mathbf{X} is on a plane $\mathbf{\Pi}$ if and only if $\mathbf{\Pi}^T\mathbf{X} = 0$.
- **Proposition 1** *Plane coordinates defined by three non collinear points can be expressed as a linear combination of the coordinates of one of these points, given by expressions (3.1) and (3.2).*

Proof:

Planes are defined by three distinct non collinear points. They are expressed by $\mathbf{U} = \begin{bmatrix} u_1 & u_2 & u_3 & u_4 \end{bmatrix}^T$, $\mathbf{V} = \begin{bmatrix} v_1 & v_2 & v_3 & v_4 \end{bmatrix}^T$ and $\mathbf{W} = \begin{bmatrix} w_1 & w_2 & w_3 & w_4 \end{bmatrix}^T$ (generating points). We search the formulation of the plane coefficients as a linear combination of one of its generating points. Consider a plane $\mathbf{\Pi}$ and define an auxiliary matrix $\mathbf{M}_{\mathbf{\Pi}} = \begin{bmatrix} \mathbf{X} & \mathbf{U} & \mathbf{V} & \mathbf{W} \end{bmatrix}$ with those three points and a generic point $\mathbf{X} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \end{bmatrix}^T$.

Since \mathbf{X} must be a linear combination of the other three points in order to belong to the plane $\mathbf{\Pi}$, the determinant of matrix $\mathbf{M}_{\mathbf{\Pi}}$ must be zero.

This gives us the expression of the plane in terms of the minors \mathbf{D}_{ijk} of matrix \mathbf{M}_{Π} (see [59]). It yields $\Pi = [\mathbf{D}_{234} \quad -\mathbf{D}_{134} \quad \mathbf{D}_{124} \quad -\mathbf{D}_{123}]^T$. After rearranging the terms, the equation can be rewritten in the form of equation (3.1) where matrix \mathbf{M} is symmetric. This equation is linear in one of the three points (\mathbf{W}) defining the plane.

$$\Pi = \mathbf{M}\mathbf{W} \quad (3.1)$$

where

$$\mathbf{M} = \begin{bmatrix} 0 & u_3v_4 - u_4v_3 & -u_2v_4 + u_4v_2 & u_2v_3 - u_3v_2 \\ -u_3v_4 + u_4v_3 & 0 & u_1v_4 - u_4v_1 & -u_1v_3 + u_3v_1 \\ u_2v_4 - u_4v_2 & -u_1v_4 + u_4v_1 & 0 & u_1v_2 - u_2v_1 \\ -u_2v_3 + u_3v_2 & u_1v_3 - u_3v_1 & -u_1v_2 + u_2v_1 & 0 \end{bmatrix} \quad (3.2)$$

■

- Planes are transformed by a point transformation \mathbf{H} in such a way that $\Pi' = \mathbf{H}^{-T}\Pi$.
- Given two planes with coordinates expressed by $\Pi_{\mathbf{A}}$ and $\Pi_{\mathbf{B}}$, the angle between them is given by its cosine expressed by equation (3.3), where Q_{∞}^* is the absolute dual quadric (see *Quadrics*).

$$\cos\theta = \frac{\Pi_{\mathbf{A}}^T \mathbf{Q}_{\infty}^* \Pi_{\mathbf{B}}}{\sqrt{(\Pi_{\mathbf{A}}^T \mathbf{Q}_{\infty}^* \Pi_{\mathbf{A}})(\Pi_{\mathbf{B}}^T \mathbf{Q}_{\infty}^* \Pi_{\mathbf{B}})}} \quad (3.3)$$

3.1.3 Lines

- Lines in space can be represented by Plücker matrices (4×4).
- A line defined by the points homogeneous \mathbf{A} and \mathbf{B} has the representation $\mathbf{L} = \mathbf{A}\mathbf{B}^T - \mathbf{B}\mathbf{A}^T$.

- Under the point transformation \mathbf{H} , the Plücker matrix transforms as $\mathbf{L}' = \mathbf{H}\mathbf{L}\mathbf{H}^T$.
- Lines can also be represented by two planes (their intersection) in Dual Plücker matrices. This representation is dual of the line Plücker matrices defined by the join of two points and is expressed by the equation $\mathbf{L}^* = \mathbf{\Pi}_A\mathbf{\Pi}_B^T - \mathbf{\Pi}_B\mathbf{\Pi}_A^T$.
- Under the point transformation \mathbf{H} the dual Plücker matrix transforms as $\mathbf{L}'^* = \mathbf{H}^{-T}\mathbf{L}\mathbf{H}^{-1}$.
- Lines have four degrees of freedom and the rank of their corresponding Plücker matrices is 2.
- The plane defined by the join of point \mathbf{X} and line \mathbf{L} is $\mathbf{\Pi} = \mathbf{L}^*\mathbf{X}$ and $\mathbf{L}^*\mathbf{X} = \mathbf{0}$ if and only if \mathbf{X} is on \mathbf{L} .
- The point defined by the intersection of the line \mathbf{L} with the plane $\mathbf{\Pi}$ is $\mathbf{X} = \mathbf{L}\mathbf{\Pi}$ and $\mathbf{L}\mathbf{\Pi} = \mathbf{0}$ if and only if \mathbf{L} is on $\mathbf{\Pi}$.
- Plücker line coordinates are defined as the six non-zero elements of the Plücker matrix. If l_{ij} is the element in its i -th line and j -th column, the line coordinates are $\ell = [l_{12} \ l_{13} \ l_{14} \ l_{23} \ l_{42} \ l_{34}]^T$.
- For a line given by the join of two points \mathbf{A} and \mathbf{B} , the Plücker line coordinates are those expressed in equation (3.4).

$$\ell = \begin{bmatrix} a_1b_2 - b_1a_2 \\ a_1b_3 - b_1a_3 \\ a_1b_4 - b_1a_4 \\ a_2b_3 - b_2a_3 \\ a_4b_2 - b_4a_2 \\ a_3b_4 - b_3a_4 \end{bmatrix} \quad (3.4)$$

- The direction of a line is its intersection with the plane at infinity $\mathbf{\Pi}_\infty = [0 \ 0 \ 0 \ 1]^T$.
- The direction of a line ℓ in P^3 is given by a 4×1 vector that depends on the Plücker line coordinates as stated in equation (3.5).

$$\mathbf{dir}(\ell) = [l_{14} \ l_{24} \ l_{34} \ 0] \quad (3.5)$$

- For the join of two points \mathbf{A} and \mathbf{B} , the direction is given by equation (3.6).

$$\mathbf{dir}(\ell_{\mathbf{AB}}) = [a_1b_4 - b_1a_4 \quad a_2b_4 - b_2a_4 \quad a_3b_4 - b_3a_4 \quad 0]^T \quad (3.6)$$

- The angle between two lines in an Euclidean 3-space is given by following equation:

$$\cos\theta = \frac{\mathbf{dir}(\ell_1)^T \mathbf{dir}(\ell_2)}{\sqrt{\mathbf{dir}(\ell_1)^T \mathbf{dir}(\ell_1)} \sqrt{\mathbf{dir}(\ell_2)^T \mathbf{dir}(\ell_2)}} \quad (3.7)$$

- **Proposition 2** *Given the lines ℓ_1 and ℓ_2 in P^3 , represented by their corresponding Plücker matrices \mathbf{L}_1 and \mathbf{L}_2 , they intersect each other if and only if $\mathbf{L}_1 \mathbf{L}_2^* \mathbf{L}_1 = \mathbf{0}$, where \mathbf{L}_2^* is the dual representation of the \mathbf{L}_2 Plücker matrix.*

Proof:

To prove that the condition is necessary we assume that the lines ℓ_1 and ℓ_2 intersect. Let us consider an **arbitrary** plane $\mathbf{\Pi}_a$. If the plane contains ℓ_1 one has $\mathbf{L}_1 \mathbf{\Pi}_a = \mathbf{0}$ and then nothing can be concluded about the matrix $\mathbf{L}_1 \mathbf{L}_2^* \mathbf{L}_1$. If the plane contains the intersection point but not line ℓ_1 nothing either can be concluded since one has $\mathbf{L}_1 \mathbf{\Pi}_a = \mathbf{X}_{1a}$

which is the point of intersection of the plane Π_a and ℓ_1 and since it is also the intersection of ℓ_1 and ℓ_2 and then is on \mathbf{L}_2 , one has $\mathbf{L}_2^* \mathbf{X}_{1a} = \mathbf{L}_2^* \mathbf{L}_1 \Pi_a = \mathbf{0}$.

However, if the plane Π_a does not contain any of the two lines nor their intersection point, $\mathbf{X}_{1a} = \mathbf{L}_1 \Pi_a$ is the intersection point of plane Π_a with line ℓ_1 . This point isn't on the line ℓ_2 and then $\Pi_b = \mathbf{L}_2^* \mathbf{X}_{1a} = \mathbf{L}_2^* \mathbf{L}_1 \Pi_a$ is the plane defined by line ℓ_2 and point \mathbf{X}_{1a} . And by hypothesis, the two lines intersect each other and then line ℓ_1 is on the plane Π_b (notice that if ℓ_1 is not on the plane Π_b the only common point with this plane is \mathbf{X}_{1a} which is not on the line ℓ_2 by definition, so one concludes that line ℓ_1 has two points on plane Π_b and consequently ℓ_1 is on Π_b). We thus have $\mathbf{L}_1 \Pi_b = \mathbf{L}_1 \mathbf{L}_2^* \mathbf{L}_1 \Pi_a = \mathbf{0}$.

Since the plane Π_a is arbitrary one thus conclude the thesis, that is $\mathbf{L}_1 \mathbf{L}_2^* \mathbf{L}_1 = \mathbf{0}$.

The counterpart should be now proved. Assume that $\mathbf{L}_1 \mathbf{L}_2^* \mathbf{L}_1 = \mathbf{0}$ and then let us try to prove that the lines intersect. Consider again an arbitrary plane Π_c not containing any of the two lines nor their intersection point. Multiplying the plane Π_c in both sides of the equation we obtain $\mathbf{L}_1 \mathbf{L}_2^* \mathbf{L}_1 \Pi_c = \mathbf{0}$, where $\mathbf{L}_1 \Pi_c$ represents the intersection point of the line ℓ_1 and the plane Π_c . Let's say \mathbf{X}_c . Since this point is not on ℓ_2 , $\mathbf{L}_2^* \mathbf{X}_c = \mathbf{L}_2^* \mathbf{L}_1 \Pi_c$ represents the plane defined by ℓ_2 and \mathbf{X}_c , say Π_d . $\mathbf{L}_1 \Pi_d = \mathbf{0}$ and then line ℓ_1 is on this plane. Since both ℓ_1 and ℓ_2 are on the same plane Π_d , they intersect each other. This proves the sufficiency of the condition. ■

- **Proposition 3** *The intersection point of two arbitrary 3-space intersecting lines ℓ_1 and ℓ_2 in P^3 is given by $\mathbf{X} = \mathbf{L}_1 \mathbf{L}_2^* \mathbf{A}$, where \mathbf{A} is an arbitrary point not belonging to the plane defined by ℓ_1 and ℓ_2 .*

Proof:

Consider plane $\Pi_{2\mathbf{A}}$ defined by an arbitrary point \mathbf{A} and ℓ_2 so that $\Pi_{2\mathbf{A}} = \mathbf{L}_2^* \mathbf{A}$. Since \mathbf{A} does not belong to the plane defined by ℓ_1 and ℓ_2 , the intersection of ℓ_1 with $\Pi_{2\mathbf{A}}$ is point \mathbf{X} , given by $\mathbf{X} = \mathbf{L}_1 \Pi_{2\mathbf{A}} = \mathbf{L}_1 \mathbf{L}_2^* \mathbf{A}$. ■

- Consider an Euclidean point \mathbf{a} and a vector \mathbf{v} and the corresponding homogeneous representation of both $\mathbf{A} = \begin{bmatrix} \mathbf{a} & 1 \end{bmatrix}^T$ and $\mathbf{V} = \begin{bmatrix} \mathbf{v} & 0 \end{bmatrix}^T$. The Plücker Matrix of line that passes in \mathbf{A} in the vector direction is given by:

$$\mathbf{L} = \mathbf{A} \cdot (\mathbf{A} + \mathbf{V})^T - (\mathbf{A} + \mathbf{V}) \cdot \mathbf{A}^T = \mathbf{A}\mathbf{V}^T - \mathbf{V}\mathbf{A}^T \quad (3.8)$$

3.1.4 Quadrics

- The point \mathbf{X} is on the quadric given by the 4×4 matrix \mathbf{Q} if and only if $\mathbf{X}^T \mathbf{Q} \mathbf{X} = 0$, where \mathbf{Q} is usually symmetric (although a non symmetric matrix \mathbf{Q} can represent the same equation).
- Quadrics are surfaces that respect the above quadratic equation whose expansion can be written as:

$$\begin{aligned} & q_{11}x_1^2 + 2q_{12}x_1x_2 + 2q_{13}x_1x_3 + 2q_{14}x_1x_4 + \\ & + q_{22}x_2^2 + 2q_{23}x_2x_3 + 2q_{24}x_2x_4 + q_{33}x_3^2 + 2q_{34}x_3x_4 + q_{44}x_4^2 = 0 \end{aligned} \quad (3.9)$$

where q_{ij} is the element (i, j) of \mathbf{Q} , considered to be symmetric.

- Under the point transformation \mathbf{H} , a point quadric transforms the the expression $\mathbf{Q}' = \mathbf{H}^{-T} \mathbf{Q} \mathbf{H}^{-1}$.
- The absolute dual quadric \mathbf{Q}_∞^* is a degenerate quadric tangent to the absolute conic Ω_∞ .

- For an Euclidean world frame, the absolute dual quadric is given by equation (3.10) where \mathbf{I}_3 is the 3×3 identity matrix.

$$\mathbf{Q}_\infty^* = \begin{bmatrix} \mathbf{I}_3 & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix} \quad (3.10)$$

- The tangent plane $\mathbf{\Pi}_N$ to a quadric in a point \mathbf{X} on the quadric surface is given by $\mathbf{\Pi}_N = \mathbf{Q}\mathbf{X}$.
- Given any non symmetric 4×4 matrix \mathbf{Q} representing a point quadric, there is a symmetric matrix \mathbf{Q}_S that represents the same point quadric. The entries of \mathbf{Q}_S are given by $q_{Sij} = \frac{q_{ij} + q_{ji}}{2}$.
- Non-ruled quadrics are quadrics whose determinant is negative. Geometrically they represent conic section surfaces such as spheroids (including ellipsoids), hyperboloids of two sheets and elliptic paraboloids. They can be represented in canonical form as function of three independent parameters: A, B and C. The following equation can be written:

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & A & B/2 \\ 0 & 0 & B/2 & -C \end{bmatrix} \quad (3.11)$$

- For each type of surface, there are some constraints that these parameters respect. For paraboloids $C = 0$ and $A = 0$. For ellipsoids $B = 0$ and for hyperboloids $A < 0$ and $C < 0$. Spheres have $A = 1$, $C + B^2/2 > 0$.
- By changing the coordinates in the x_3 axis for hyperboloids and spheres, the additional constraint $B = 0$ can be applied. Thus, it can be re-

garded that the parameter B is non zero only for paraboloids and zero for all other non-ruled surface types.

- In Euclidean coordinates, the normal vector \mathbf{n} to a quadric \mathbf{Q} at a point \mathbf{X} is given by the first three coordinates of the normal plane coordinates $\mathbf{\Pi}_N = \mathbf{Q}\mathbf{X}$. $n_i = \mathbf{\Pi}_{N_i}$, with $i \in \{1, 2, 3\}$. It is sometimes useful to normalize the normal vector so that it becomes unitary.
- **Proposition 4** *The normal line to the quadric \mathbf{Q} at point \mathbf{X} is given by the Plücker matrix expressed as $\mathbf{L}_N = \mathbf{X}\mathbf{X}^T\mathbf{Q}^T\mathbf{Q}_\infty^* - \mathbf{Q}_\infty^*\mathbf{Q}\mathbf{X}\mathbf{X}^T$, where \mathbf{Q}_∞^* is the dual absolute quadric.*

Proof:

The tangent plane to the quadric through \mathbf{X} is given by $\mathbf{\Pi}_N = \mathbf{Q}\mathbf{X}$ and the direction of this plane is given by $\mathbf{dir}(\mathbf{\Pi}_N) = \mathbf{Q}_\infty^*\mathbf{\Pi}_N = \mathbf{Q}_\infty^*\mathbf{Q}\mathbf{X}$. Since the direction of a plane also represents the intersection of its normal line with the plane at infinity [121], the normal line is the join of the points \mathbf{X} and $\mathbf{dir}(\mathbf{\Pi}_N)$, given by $\mathbf{L}_N = \mathbf{X}\mathbf{dir}(\mathbf{\Pi}_N)^T - \mathbf{dir}(\mathbf{\Pi}_N)\mathbf{X}^T$ or $\mathbf{L}_N = \mathbf{X}\mathbf{X}^T\mathbf{Q}^T\mathbf{Q}_\infty^* - \mathbf{Q}_\infty^*\mathbf{Q}\mathbf{X}\mathbf{X}^T$. ■

3.2 Some Geometric Relations in 3D Space

In this section the points are expressed in Euclidean coordinates.

The distance of a point \mathbf{X}_0 to a line ℓ in 3D space defined by two points \mathbf{X}_1 and \mathbf{X}_2 is given by the following equation [136]:

$$distance = \frac{\|(\mathbf{X}_2 - \mathbf{X}_1) \times (\mathbf{X}_1 - \mathbf{X}_0)\|}{\|\mathbf{X}_2 - \mathbf{X}_1\|} \quad (3.12)$$

The point in line ℓ that is closer to the point \mathbf{X}_0 is given by the following expression:

$$\mathbf{X}_{0\ell} = \mathbf{X}_1 - \frac{(\mathbf{X}_1 - \mathbf{X}_0) \cdot (\mathbf{X}_2 - \mathbf{X}_1)}{\|\mathbf{X}_2 - \mathbf{X}_1\|^2} \cdot (\mathbf{X}_2 - \mathbf{X}_1) \quad (3.13)$$

Consider now two skew lines in space, each of them defined by two points: $(\mathbf{X}_1, \mathbf{X}_2)$ and $(\mathbf{X}_3, \mathbf{X}_4)$. The distance between these two lines is given by the expression:

$$distance = \frac{\|(\mathbf{X}_3 - \mathbf{X}_1) \cdot ((\mathbf{X}_2 - \mathbf{X}_1) \times (\mathbf{X}_4 - \mathbf{X}_3))\|}{\|(\mathbf{X}_2 - \mathbf{X}_1) \times (\mathbf{X}_4 - \mathbf{X}_3)\|} \quad (3.14)$$

The common orthogonal line to both these lines is defined by the two intersecting points \mathbf{I}_1 and \mathbf{I}_2 . As stated in [39], the minimum distance between two skew lines in R^n is achieved for points in lines ℓ_1 and ℓ_2 such that:

$$\begin{cases} \mathbf{I}_1 = \mathbf{X}_1 + t \cdot (\mathbf{X}_2 - \mathbf{X}_1) = \mathbf{X}_1 + t \cdot \mathbf{V}_{12} \\ \mathbf{I}_2 = \mathbf{X}_3 + s \cdot (\mathbf{X}_4 - \mathbf{X}_3) = \mathbf{X}_3 + s \cdot \mathbf{V}_{34} \end{cases} \quad (3.15)$$

where \mathbf{V}_{12} and \mathbf{V}_{34} are the line direction vectors. The distance equation has a single solution if and only if these vectors are not parallel, that is, if the following condition is met:

$$\det \begin{bmatrix} \|\mathbf{V}_{12}\|^2 & -\mathbf{V}_{12} \cdot \mathbf{V}_{34} \\ -\mathbf{V}_{12} \cdot \mathbf{V}_{34} & \|\mathbf{V}_{34}\|^2 \end{bmatrix} \neq 0 \quad (3.16)$$

The minimum distance between two skew lines then occurs to:

$$\begin{cases} s = \frac{\|\mathbf{V}_{12}\|^2(\mathbf{V}_{34} \cdot \mathbf{X}_3 - \mathbf{V}_{34} \cdot \mathbf{X}_1) + (\mathbf{V}_{12} \cdot \mathbf{X}_1 - \mathbf{V}_{12} \cdot \mathbf{X}_3)(\mathbf{V}_{12} \cdot \mathbf{V}_{34})}{(\mathbf{V}_{12} \cdot \mathbf{V}_{34})^2 - \|\mathbf{V}_{12}\|^2 \|\mathbf{V}_{34}\|^2} \\ t = \frac{(\mathbf{V}_{12} \cdot \mathbf{V}_{34})s - (\mathbf{V}_{12} \cdot \mathbf{X}_1 - \mathbf{V}_{12} \cdot \mathbf{X}_3)}{\|\mathbf{V}_{12}\|^2} \end{cases} \quad (3.17)$$

3.3 Euler Angles and Rotation Matrices

The calculation of a rotation matrix in 3D space, from the corresponding Euler angles around the coordinate axes is straightforward. However, for the reciprocal one has to assume a sequence of three rotations (three different axes or two different axes by interchanging their order). Actually, this process is not injective and then, for a given rotation matrix, there are different sets of possible generating Euler angles. This section reviews an algorithm to calculate all different sets of angles that correspond to a given rotation matrix [118].

Consider the rotation matrix represented by $\mathbf{Rot} = [Rot_{ij}]$, with $(i, j) \in \{1, 2, 3\}$ correspondent to the Euler angles ψ , θ and ϕ performing rotations about the x_1 , x_2 and x_3 coordinate axis respectively.

If $Rot_{31} \neq \pm 1$, then we have the following two sets of Euler angles:

$$\left\{ \begin{array}{l} \theta_1 = -asin(Rot_{31}) \\ \psi_1 = atan2\left(\frac{Rot_{32}}{\cos\theta_1}, \frac{Rot_{33}}{\cos\theta_1}\right) \\ \phi_1 = atan2\left(\frac{Rot_{21}}{\cos\theta_1}, \frac{Rot_{11}}{\cos\theta_1}\right) \end{array} \right. \quad \left\{ \begin{array}{l} \theta_2 = \pi - asin(Rot_{31}) \\ \psi_2 = atan2\left(\frac{Rot_{32}}{\cos\theta_2}, \frac{Rot_{33}}{\cos\theta_2}\right) \\ \phi_2 = atan2\left(\frac{Rot_{21}}{\cos\theta_2}, \frac{Rot_{11}}{\cos\theta_2}\right) \end{array} \right. \quad (3.18)$$

On the other and, if $Rot_{31} = \pm 1$, the two sets of Euler angles are:

$$\left\{ \begin{array}{l} \text{if } Rot_{31} = -1 \\ \theta_3 = \pi/2 \\ \phi_3 = \text{anything, can be set to } 0 \\ \psi_3 = \phi_3 + atan2(Rot_{12}, Rot_{13}) \end{array} \right. \quad \left\{ \begin{array}{l} \text{if } Rot_{31} = 1 \\ \theta_4 = -\pi/2 \\ \phi_4 = \text{anything, can be set to } 0 \\ \psi_4 = -\phi_4 + atan2(Rot_{12}, Rot_{13}) \end{array} \right. \quad (3.19)$$

where $asin$ is the arc-sine and $atan2$ is the arc-tangent.

These expressions for the sets of Euler angles that corresponds to a given rotation matrix assume that the matrix \mathbf{Rot} is a rotation matrix. A rotation

matrix respects the following constraints: its determinant is unitary (for a proper rotation), the norm of each column and row is unitary, it is orthogonal ($\mathbf{Rot}^T = \mathbf{Rot}^{-1}$) and in what concerns the eigenvalues, they satisfy one of the following conditions:

- All eigenvalues are equal to 1.
- One eigenvalue is 1 and the other two are equal to -1.
- One eigenvalue is 1 and the other two are complex conjugate of the form $e^{\pm i\theta}$.

Another useful property of rotation matrices is that all their singular values are equal to 1.

However, often the transformation matrix available is not a perfect rotation matrix, since it does not respect the above constraints. The rotation matrices computed by minimization algorithms usually have this problem. To recover the Euler angles as accurately as possible, the rotation matrix should be previously manipulated so that the determinant is unitary and the corresponding eigenvalues met one of the conditions. Additionally, all its singular values must be equal to 1. This is easily accomplished by performing a singular value decomposition.

Consider the rotation matrix \mathbf{Rot} and its singular value decomposition such that $\mathbf{Rot} = \mathbf{U} \cdot \mathbf{S} \cdot \mathbf{V}^T$, where \mathbf{U} and \mathbf{V} are orthogonal square matrices and \mathbf{S} is a diagonal square matrix with the \mathbf{Rot} singular values in its diagonal. In a perfect rotation matrix, all these singular values are equal to 1 so matrix \mathbf{S} is the identity matrix. The new corrected rotation matrix can then be calculated by multiplying the orthogonal matrices \mathbf{U} and \mathbf{V}^T , such that $\mathbf{Rot}_{\text{orth}} = \mathbf{UV}^T$ [114, 115]. Matrix $\mathbf{Rot}_{\text{orth}}$ is now an orthogonal matrix and can be used to compute Euler angles with higher accuracy by computing equation (3.18) or (3.19).

3.4 Quaternions

Rotations can also be expressed by quaternions, formalized by the mathematician Sir Hamilton in the 19th century. When compared to Euler angles and rotation matrices, quaternions are sometimes used since the number of parameters to estimate is smaller (4 parameters using quaternions and 9 using rotation matrices) and also because, as pointed out by Triggs et al. [132] generally quaternions present better estimating properties. Usually Euler angles present some numerical problems and furthermore quaternions have a behavior closer to linear.

This section reviews the basics of quaternion algebra. For more details regarding the algebraic properties of quaternions see, for instance [135].

Quaternions are quadruples of real numbers defined as $\mathbf{q} = [s, (q_1, q_2, q_3)] = [s, \mathbf{v}]$.

The addition and multiplication of two quaternions $\mathbf{q}_a = [s_a, \mathbf{v}_a]$ and $\mathbf{q}_b = [s_b, \mathbf{v}_b]$ are given by:

$$\mathbf{q}_a + \mathbf{q}_b = [s_a + s_b, \mathbf{v}_a + \mathbf{v}_b] \quad (3.20)$$

$$\mathbf{q}_a \mathbf{q}_b = [s_a s_b - \mathbf{v}_a^T \mathbf{v}_b, s_a \mathbf{v}_b + s_b \mathbf{v}_a + \mathbf{v}_a \times \mathbf{v}_b] \quad (3.21)$$

and the inverse of a quaternion is $\mathbf{q}^{-1} = \|\mathbf{q}\|^{-2} [s, -\mathbf{v}]$.

Quaternions are used to represent position vectors in 3D space such that a point $\mathbf{X} = [X_1, X_2, X_3]^T$ is represented by the quaternion $\mathbf{X}_q = [0, (X_1, X_2, X_3)]^T$.

Quaternions can also be used to perform rotations about an axis. Suppose you have a point \mathbf{X}_q in space and want to perform a rotation about an arbitrary axis. This rotation can be decomposed into three elementary rotations about the Cartesian orthogonal axis OX_1 , OX_2 and OX_3 by, respectively, θ_{pitch} , θ_{yaw} and θ_{roll} . The equations for the unitary rotation quaternion that

transforms point \mathbf{X}_q are the following:

$$\begin{aligned}
s &= C(\theta_y)C(\theta_p)C(\theta_r) + S(\theta_y)S(\theta_p)S(\theta_r) \\
q_1 &= C(\theta_y)S(\theta_p)C(\theta_r) + S(\theta_y)C(\theta_p)S(\theta_r) \\
q_2 &= S(\theta_y)C(\theta_p)C(\theta_r) - C(\theta_y)S(\theta_p)S(\theta_r) \\
q_3 &= C(\theta_y)C(\theta_p)S(\theta_r) - S(\theta_y)S(\theta_p)C(\theta_r)
\end{aligned} \tag{3.22}$$

where $\theta_y = \theta_{yaw}/2$, $\theta_p = \theta_{pitch}/2$ and $\theta_r = \theta_{roll}/2$. $C()$ and $S()$ represent respectively the cosine and sine of the argument angle.

The equation that transforms point \mathbf{X}_q into point \mathbf{X}'_q is:

$$\mathbf{X}'_q = \mathbf{q}_{rot} \mathbf{X}_q \mathbf{q}_{rot}^{-1} \tag{3.23}$$

where the elementary rotations are performed in the pre defined order given by $\mathbf{q}_{rot} = \mathbf{q}_{yaw} \mathbf{q}_{pitch} \mathbf{q}_{roll}$.

The relation between quaternions and general rotation matrices is given by:

$$\mathbf{Rot} = \begin{bmatrix} 1 - 2(q_2^2 + q_3^2) & 2(q_1q_2 - sq_3) & 2(q_1q_3 + sq_2) \\ 2(q_1q_2 + sq_3) & 1 - 2(q_1^2 + q_3^2) & 2(q_2q_3 + sq_1) \\ 2(q_1q_3 + sq_3) & 2(q_2q_3 + sq_1) & 1 - 2(q_1^2 + q_2^2) \end{bmatrix} \tag{3.24}$$

such that the relationship $\mathbf{X}' = \mathbf{TX}$ is equivalent to expression (3.23).

If homogeneous coordinates are used, the 4×4 transformation matrix includes the rotation matrix \mathbf{Rot} and a translation vector as:

$$\mathbf{T} = \begin{bmatrix} \mathbf{Rot} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \tag{3.25}$$

3.5 Intersection of Two Quadrics

The parameterization of the intersection of two quadrics is a topic still studied. As it is used in the development of this dissertation, we here review one recent method for the computation of this intersection. Dupont et al. and Lazard et al. [32, 33, 81] presented a method to intersect two general quadrics. The solution is a quartic curve in space, parameterized by one scalar parameter.

Before we present the intersection algorithm, some notions about quadrics and their intersection are first reviewed.

As stated by [117] (Semple and Kneebone, 1959) the intersection of two quadrics \mathbf{Q} and \mathbf{S} is the quartic curve that is also common to all the quadrics in the pencil defined by $\lambda\mathbf{Q} + \mathbf{S}$. Both Levin's original method and the Lazard et al. are based in this property of the intersection of quadrics.

The inertia of a quadric is defined by the number of its positive and negative eigenvalues. Sylvester's Inertia Law states that the inertia of a quadric is invariant to a real projective transformation [80]. The inertia of \mathbf{S} is represented by $\sigma_S = (\max(\sigma^+, \sigma^-), \min(\sigma^+, \sigma^-))$, where σ^+ and σ^- are the number of positive and negative eigenvalues respectively. Table 3.1 presents all possible inertias, their physical realization and the corresponding canonical form.

Since this method is, in our work, used to compute the intersection between a mirror surface, expressed by a non-ruled quadric, and another general quadric, we assume that one of the quadrics has inertia $(3, 1)$ (say quadric \mathbf{Q}) and make no assumptions about the other one (say quadric \mathbf{S}). The algorithm is now reviewed.

Table 3.1: Euclidean type of a quadric \mathbf{Q} in terms of its inertia.

Inertia of \mathbf{S}	Affine reduced canonical form	Euclidean type of \mathbf{S}
(4, 0)	$x^2 + y^2 + z^2 + 1 = 0$	\emptyset (imaginary ellipsoid)
(3, 1)	$x^2 + y^2 + z^2 - 1 = 0$ $x^2 + y^2 - z^2 + 1 = 0$ $x^2 + y^2 + z = 0$	ellipsoid hyperboloid of two sheets elliptic paraboloid
(3, 0)	$x^2 + y^2 + z^2 = 0, x^2 + y^2 + 1 = 0$	point, \emptyset (imaginary elliptic cylinder)
(2, 2)	$x^2 + y^2 - z^2 - 1 = 0, x^2 - y^2 + z = 0$	hyperboloid of one sheet hyperbolic paraboloid
(2, 1)	$x^2 + y^2 - z^2 = 0, x^2 + y^2 - 1 = 0$ $x^2 - y^2 + 1 = 0, x^2 + y = 0$	elliptic cylinder, hyperbolic cylinder cone, parabolic cylinder
(2, 0)	$x^2 + y^2 = 0, x^2 + 1 = 0$	line, \emptyset (imaginary parallel planes)
(1, 1)	$x^2 - y^2 = 0, x^2 - 1 = 0, x = 0$	parallel planes, simple plane intersecting planes
(1, 0)	$x^2 = 0$	double plane, \emptyset

Algorithm for the intersection of two quadrics:

- **Step 1** - Find a quadric \mathbf{R} in the pencil $\mathbf{R}(\lambda) = \lambda\mathbf{Q} - \mathbf{S}$ such that $\det(\mathbf{R}) > 0$ or otherwise $\det(\mathbf{R}) = 0$. If there is no such quadric, the intersection is reduced to two points and if the inertia of \mathbf{R} is $(4,0)$, the intersection is an empty set.

Notice that as the quadric \mathbf{Q} represents the mirror which is a non-ruled quadric (an ellipsoid, hyperboloid of two sheets or paraboloid), its determinant is always negative. Therefore, we state $\mathbf{R} = \mathbf{S}$.

Levin [82, 83] has proved that there is at least one ruled quadric in any pencil of quadrics. Ruled quadrics (those quadrics whose inertia is different from $(3,1)$) have positive or zero determinant so this step is always solvable by first finding the zeros of the determinant of the upper 3×3 submatrix of $\mathbf{R}(\lambda)$, a cubic equation. See implementation details in [32].

- **Step 2** - If the inertia of \mathbf{R} is different from $(2,2)$, apply Gauss reduction of quadratic forms to \mathbf{R} . Compute the linear transformation matrix \mathbf{P} such that $\mathbf{P}^T\mathbf{R}\mathbf{P}$ is diagonal. The quadric can then be reduced to its canonical form. By Sylvester's Inertia Law [80], the inertia of the reduced form quadric $\mathbf{R}_c = \mathbf{P}^T\mathbf{R}\mathbf{P}$ is the same as the inertia of \mathbf{R} . Choose the appropriate parameterization of the quadric in table 3.2.

If the inertia of \mathbf{R} is $(2,2)$, choose a quadric \mathbf{R}' from the pencil of quadrics that passes through a point close enough from \mathbf{R} and rename \mathbf{R}' by \mathbf{R} . Compute a matrix \mathbf{P} such that $\mathbf{R}_c = \mathbf{P}^T\mathbf{R}\mathbf{P}$ is a diagonal matrix with the elements in the diagonal equal to $(1, 1, -1, -\delta)$, where δ is a positive rational number. See [32] for details.

Table 3.2: Optimal parameterization of the projective quadrics \mathbf{R}_c of inertia different from $(3, 1)$. The parameterized point \mathbf{X}_c satisfy the quadric equation in its canonical form: $\mathbf{X}_c^T \mathbf{R}_c \mathbf{X}_c = 0$. See [32] for the details.

Inertia of \mathbf{R}_c	Canonical equation ($a, b, c, d > 0$)	Parameterization $\mathbf{X}_c = [x \ y \ z \ w]$
(4, 0)	$ax^2 + by^2 + cz^2 + dw^2 = 0$	\mathbf{R}_c is the empty set \emptyset
(3, 0)	$ax^2 + by^2 + cz^2 = 0$	\mathbf{R}_c is a point
(2, 2)	$ax^2 + by^2 - cz^2 - dw^2 = 0$	$\mathbf{X}_c = \left[\frac{ut+avs}{a}, \frac{us-bvt}{b}, \frac{ut-avs}{\sqrt{ac}}, \frac{us+bvt}{\sqrt{bd}} \right]$ with $(u, v), (s, t) \in P^1$
(2, 1)	$ax^2 + by^2 - cz^2 = 0$	$\mathbf{X}_c = \left[uv, \frac{u^2-abv^2}{2b}, \frac{u^2+abv^2}{2\sqrt{bc}}, s \right]$ with $(u, s, v) \in P^2$
(2, 0)	$ax^2 + by^2 = 0$	$\mathbf{X}_c = [0, 0, u, v]$ with $(u, v) \in P^1$
(1, 1)	$ax^2 - by^2 = 0$	$\mathbf{X}_{c1} = \left[u, \frac{\sqrt{ab}}{b}u, v, s \right], \mathbf{X}_{c2} = \left[u, -\frac{\sqrt{ab}}{b}u, v, s \right]$ with $(u, v, s) \in P^2$ (quadric in two branches)
(1, 0)	$ax^2 = 0$	$\mathbf{X}_c = [0, u, v, s]$ with $(u, v, s) \in P^2$

In the local canonical frame $\mathbf{R}_c = \mathbf{P}^T \mathbf{R} \mathbf{P}$ represents the quadric \mathbf{R} in one of the canonical representations parameterized by any \mathbf{X}_c of table 3.2. Compute then the parameterization $\mathbf{X}_o = \mathbf{P} \mathbf{X}_c$ of \mathbf{R} in the original coordinate system.

In order to reduce the number of parameters of the parameterization of \mathbf{X}_o and since they are defined up to a scale factor, in P^1 or P^2 , it is possible to state one or two parameters to a fixed scale factor. We thus state $v = 1$ (and possibly also $t = 1$ for the inertia $(2, 2)$ case) in P^1 and we state $s = 1$ in P^2 parameterizations. Notice that eliminating the scale factor dependency, the parameterization is now written in only one or two parameters.

- **Step 3** - Consider the equation $\mathbf{X}_o^T \mathbf{Q} \mathbf{X}_o = 0$. By expanding it, one obtains an equation of degree two in at least one of the parameters u , v or s . Solve this equation for one of them and compute the domain of the solution. Then substitute this parameter as a function of the other in \mathbf{X}_o . The equation obtained is the parameterization of the intersection of \mathbf{Q} and \mathbf{S} .

This algorithm is optimal in terms of number of irrationals in the solution and is near-optimal in terms of the size of the solution. As pointed out by their authors, the introduction of irrationals can highly perturb the accuracy of the computation.

Part II

MODELS AND METHODS

Chapter 4

Geometry of Catadioptric Image Formation

The aim of this chapter is the introduction to the framework and the analysis of the image formation geometry. Some new results are presented on this topic with a restriction on the projection model of noncentral catadioptric vision systems with quadric mirrors.

Due to useful properties, homogeneous coordinates are used to express geometric entities and their relations. Quadrics are expressed by the usual 4×4 symmetric matrix $\mathbf{Q} = \{q_{ij}\}$ and lines are defined by the corresponding Plücker matrices.

Our noncentral catadioptric vision systems are composed by a pinhole camera and a quadric surface mirror. The position of the camera in relation to the mirror is not restricted to the symmetry axis.

4.1 Image Formation and the Projection Model

In a catadioptric vision system the light rays are reflected by a specular mirror in the direction of a light sensor, usually a CCD. Light rays are called

incident before they interact with the specular surface and reflected after this interaction. In this chapter we are interested in the analysis of the reflection geometry.

The geometry of the projection of the reflection point is well known and depends on the camera. The problem is how to find the reflection point where the light travelling from an arbitrary 3D point is reflected to the image. This mapping between the 3D points and the image 2D points is the projection model searched for.

There is a key distinction between central systems and non central systems. Central projection happens for those systems whose geometry is such that all incident light rays intersect each other at a unique point, called effective viewpoint, irrespective of the location of the point reflected to the image sensor. The image is virtually taken from this viewpoint. When this unique point doesn't exist and the incident light rays are skew rather than concurrent, the system as well as the projection are, by contrary, called non-central.

For central projection systems the projection model is well studied and closed form expressions are provided. However, for noncentral projections, the non existence of an effective viewpoint difficult the geometry of image formation and as long as the author is aware of there is no closed form explicit solution for the image point where a given 3D point is imaged. The principles of physics are then used to understand and formulate the projection model of noncentral catadioptric vision systems.

There are two principles that describe the reflection process. On one hand, the Snell's Law states that the reflection point is the surface point whose normal vector is the bisector of the incident and reflected light rays (the reflected ray passes through the known optical center of the pinhole camera). It also states that this normal vector to the surface is within the plane defined by the optical center, the point to be projected and the reflection

point (see [60]).

On the other hand, by the laws of the Optical Geometry, it is known that the reflection point is the one that makes the light path to be the quickest one. This principle is called Fermat Principle and its first formulation is dated from 1657, based on the ancient variational principle by Hero of Alexandria (somewhere between 150 BC and 250 AD) - see [60]. Since these distances are small and no perturbation happens in the space-time, the quickest path is also the shortest one and so the total path can be minimized to achieve the reflection point.

Both Snell's Law and Fermat principle are sufficient, each one **per si**, to find the reflection point. The problem is that the constraints are not explicit in the image coordinates and to solve them it is necessary to solve a multidimensional nonlinear system of equations. The dimension of the problem depends on the formulation. This problem is not difficult to solve since the expressions are well behaved for the majority of the vision systems but it is slow and computationally intense.

Assume that the camera center, the quadric mirror and the 3D point to project into image are known in camera coordinates. Also consider Cartesian coordinates for the computation of the reflection point on the specular surface by Snell's Law and Fermat Principle.

4.1.1 Snell's Law

By the Snell's Law, the incident and reflected light rays are at an equal angle in relation to the normal direction to the mirror surface at the reflection point \mathbf{r} . Furthermore, the same reflection point \mathbf{r} , the camera center \mathbf{c} and the point to project \mathbf{p} define a plane that contains the normal vector.

Figure 4.1 shows the reflection process where \mathbf{v}_i is the incident light ray, \mathbf{v}_r is the reflected light ray and \mathbf{n} is the normal vector to the mirror surface. The reflection law, in Euclidean coordinates, is then given by the equation:

$$\mathbf{v}_r = \mathbf{v}_i - 2(\mathbf{v}_i \cdot \mathbf{n})\mathbf{n} \quad (4.1)$$

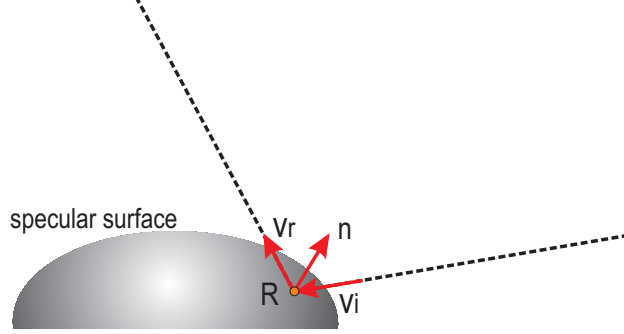


Figure 4.1: Specular reflection

To express the reflected ray \mathbf{v}_r we use an additional constraint to the equation, that all reflected light rays pass through the optical center of the camera \mathbf{c} .

The expression of the 3×1 normal vector is given by equation:

$$\mathbf{n} = \frac{\frac{\partial \mathbf{x}}{\partial x_1} \times \frac{\partial \mathbf{x}}{\partial x_2}}{\left\| \frac{\partial \mathbf{x}}{\partial x_1} \times \frac{\partial \mathbf{x}}{\partial x_2} \right\|} \quad (4.2)$$

where $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}$ is the generic point on the mirror surface whose coordinates are related by the quadratic surface equation (the quadric mirror equation (3.9), where for Cartesian coordinates $x_4 = 1$). The quadric surface equation is then: $q_{11}x_1^2 + 2q_{12}x_1x_2 + 2q_{13}x_1x_3 + 2q_{14}x_1 + q_{22}x_2^2 + 2q_{23}x_2x_3 + 2q_{24}x_2 + q_{33}x_3^2 + 2q_{34}x_3 + q_{44} = 0$.

Since we know the camera center \mathbf{c} and the 3D point \mathbf{p} , equation (4.1) can be used in the form:

$$\frac{\mathbf{c} - \mathbf{r}}{\|\mathbf{c} - \mathbf{r}\|} = \frac{\mathbf{r} - \mathbf{p}}{\|\mathbf{r} - \mathbf{p}\|} - 2 \left(\frac{\mathbf{r} - \mathbf{p}}{\|\mathbf{r} - \mathbf{p}\|} \cdot \mathbf{n} \right) \mathbf{n} \quad (4.3)$$

This equation is not explicit to the reflection point \mathbf{r} . It is easy to solve for \mathbf{r} but computationally hard due to all nonlinearities introduced by the norms (all incident and reflected light rays have unit norm) and by the normal vector.

4.1.2 Fermat Principle

The reflection point can also be calculated using the Fermat principle. This principle states that the light always takes the quickest path. So the reflection point is the one that minimizes the distance between the 3D point \mathbf{p} and the camera center \mathbf{c} . Notice that for the order of magnitude of these systems, no perturbation in the space-time exists and so the quickest path is also the shortest one.

Since we also know the quadric mirror parameters, it is possible to express one of the coordinates as function of the other two. We opt to express x_3 in relation to x_1 and x_2 . This is done to incorporate the mirror restriction in the equation of Fermat principle. The coordinates of the reflection point are then expressed in Cartesian coordinates by $\mathbf{r} = [r_1 \ r_2 \ r_3]$, where the third coordinate is given by the following equation:

$$r_3 = -\frac{1}{c} (q_{13}r_1 + q_{23}r_2 + q_{34}) \pm \frac{1}{2q_{33}} \sqrt{F_3} \quad (4.4)$$

where

$$F_3 = (2q_{13}r_1 + 2q_{23}r_2 + 2q_{34}) - 4q_{33} (q_{11}r_1^2 + q_{22}r_2^2 + 2q_{12}r_1r_2 + 2q_{14}r_1 + 2q_{24}r_2 + q_{44}) \quad (4.5)$$

and the appropriate root must be chosen.

The distances between \mathbf{r} and \mathbf{c} and between \mathbf{r} and \mathbf{p} can now be calculated and their sum minimized. The total distance is then given by:

$$d_{light} = \sqrt{(r_1 - c_1)^2 + (r_2 - c_2)^2 + (r_3 - c_3)^2} + \sqrt{(r_1 - p_1)^2 + (r_2 - p_2)^2 + (r_3 - p_3)^2} \quad (4.6)$$

where r_3 is given by expression (4.5).

Expression (4.6) can be analytically minimized by any known method. The expressions obtained are nonlinear and implicit in the coordinates of the reflection point. It still needs a nonlinear minimization method to compute numerically the solution.

4.2 A New Projection Model

From equation (4.3) and (4.6) we see that both the Snell's Law and the Fermat Principle solve the problem of the reflection point. The solution is however, implicit, non linear, often unstable and computationally demanding.

In this section we present a projection model that can be applied to noncentral catadioptric vision systems composed by a quadric surface mirror and a perspective projection camera. The camera intrinsic parameters, the quadric and the pose of the camera in relation to the mirror are assumed to be known. Homogeneous coordinates are used rather than Cartesian. These issues were partially addressed by Gonçalves and Araújo in [50, 51].

Consider then the camera whose optical center is the point \mathbf{C} and the intrinsic parameters matrix is the matrix \mathbf{K} . The mirror surface is given by a quadric \mathbf{Q} and is positioned freely in relation to the camera. The 3D world point \mathbf{P} is imaged by the camera and its reflection point over the mirror surface is the point \mathbf{R} . Figure 4.2 shows the reflection process and the notations adopted.

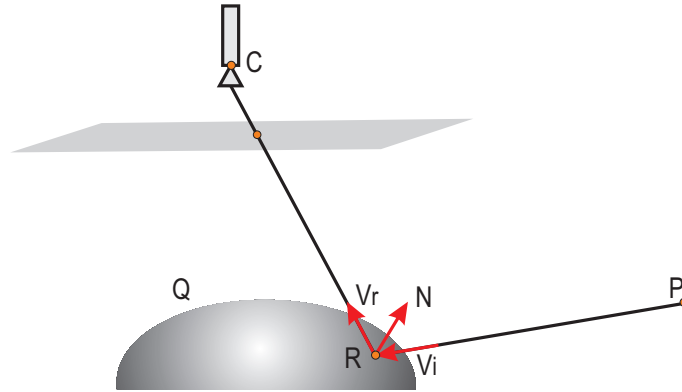


Figure 4.2: The light rays reflection and imaging in a catadioptric vision system.

4.2.1 Restrictions imposed on the reflection point

\mathbf{R} is the reflection point on the mirror surface that projects the 3D point \mathbf{P} into the image plane passing through the camera center \mathbf{C} . For such point the following three restrictions must be imposed:

1. $\mathbf{R}^T \mathbf{Q} \mathbf{R} = 0 \longrightarrow$ the point is on the quadric of the mirror surface.
2. $\mathbf{R}^T \mathbf{S} \mathbf{R} = 0 \longrightarrow$ the point is on the quadric given by $\mathbf{S} = \mathbf{M}^T \mathbf{Q}_\infty^* \mathbf{Q}$ (proposition 5).

Proposition 5 *The reflection point \mathbf{R} of a catadioptric system with quadric mirror \mathbf{Q} , reflecting a 3D world point \mathbf{P} to the camera optical center \mathbf{C} , is on the quadric surface \mathbf{S} , given by $\mathbf{S} = \mathbf{M}^T \mathbf{Q}_\infty^* \mathbf{Q}$, where \mathbf{Q}_∞^* is the absolute dual quadric, the 4×4 matrix \mathbf{M} is given*

by expression (3.2) and the plane Π_B is defined by the 3D world point \mathbf{P} , the camera optical center \mathbf{C} and the reflection point \mathbf{R} itself. The reflection point \mathbf{R} is such that $\Pi_B = \mathbf{MR}$.

Proof:

Let us consider two concurrent planes: Π_A and Π_B . Π_A is the tangent plane to the quadric \mathbf{Q} at the reflection point \mathbf{R} . Its representation is given by $\Pi_A = \mathbf{QR}$.

The plane Π_B is the plane defined by three points: the camera optical center \mathbf{C} , the 3D point \mathbf{P} and the reflection point \mathbf{R} on the mirror surface. Using equation (3.1) the plane coordinates vector can be defined by a linear equation in the reflected point \mathbf{R} as stated by $\Pi_B = \mathbf{M}(\mathbf{P}, \mathbf{C}) \cdot \mathbf{R} = \mathbf{MR}$ (see equation (3.2)).

Since the normal to the quadric is perpendicular to the tangent plane and must be on the plane defined by the three points \mathbf{C} , \mathbf{P} and \mathbf{R} , then the two planes, Π_A and Π_B , must be perpendicular. The angle between two planes is given by equation (3.3) and since we admit an Euclidean space, the absolute dual quadric for Euclidean transformations is given by expression (3.10).

Since $\theta = \pi/2$ and substituting equations of the planes Π_A and Π_B into equation (3.3) it yields equation (4.7) which restricts the point \mathbf{R} to be on a quadric surface given by $\mathbf{S} = \mathbf{M}^T \mathbf{Q}_\infty^* \mathbf{Q}$.

$$\begin{aligned} \Pi_A^T \mathbf{Q}_\infty^* \Pi_B = 0 &\Leftrightarrow \mathbf{R}^T \mathbf{Q}^T \mathbf{Q}_\infty^* \mathbf{MR} = 0 \Leftrightarrow \\ &\Leftrightarrow \mathbf{R}^T \mathbf{M}^T \mathbf{Q}_\infty^* \mathbf{QR} = 0 \end{aligned} \quad (4.7)$$

Notice that matrix \mathbf{S} is not symmetric as the generic quadric matrix. However, without loss of generality, matrix \mathbf{S} can be substituted by

another matrix whose entries are related by $S_{ij} \leftarrow 0.5S_{ij} + 0.5S_{ji}$. With this change the quadric remains the same and its representing matrix becomes symmetric. ■

3. The incidence and reflected angles are equal.

The normal line $\ell_{\mathbf{N}}$ is the bisector of the angle between incident and reflection lines. Using equation (3.7), after some simplifications one obtains:

$$\frac{\mathbf{dir}(\ell_{\mathbf{RC}})^T \mathbf{dir}(\ell_{\mathbf{N}})}{\sqrt{\mathbf{dir}(\ell_{\mathbf{RC}})^T \mathbf{dir}(\ell_{\mathbf{RC}})}} = \frac{\mathbf{dir}(\ell_{\mathbf{PR}})^T \mathbf{dir}(\ell_{\mathbf{N}})}{\sqrt{\mathbf{dir}(\ell_{\mathbf{PR}})^T \mathbf{dir}(\ell_{\mathbf{PR}})}} \quad (4.8)$$

The directions $\ell_{\mathbf{RC}}$ and $\ell_{\mathbf{PR}}$ are defined by the join of two points using equation (3.6) and $\ell_{\mathbf{N}}$ is computed using proposition 4.

4.2.2 Computing the reflection point \mathbf{R}

Given the three constraints imposed to the reflection point \mathbf{R} , the problem is now how to find that point. Its explicit closed form computation is however still not possible. The first and second constraints are much similar since they restrict the point \mathbf{R} to be on quadric \mathbf{Q} (constraint (1)) and to be also on quadric \mathbf{S} (constraint (2)). This is the problem of finding the intersection of those two quadrics (a quartic in space). Since the third restriction constrains the point so that the incident and reflection angles are equal, point \mathbf{R} must be located on the intersection curve.

The general method for computing an explicit parametric representation of the intersection between two quadrics is due to Joshua Levin [82,83]. However, the parametric representation of this method is hard to compute and is less reliable due to the high number of irrational numbers needed. Dupont et

al. [32,33] presented a modification of the Levin method to intersect quadrics with optimal number of irrationals, pointing out that this alternative method is much more accurate than the original one. This method is reviewed in section 3.5.

The parametric curve given by the intersection algorithm is a function of only one parameter, say λ . Let us represent the parameterized curve by the 4×1 vector $\mathbf{X}(\lambda)$. Although nonlinear, the curve can be searched for the point where incident and reflected angles are equal, that is, where equation (4.8) holds. Let us call λ_0 to the value of the parameter that solves equation (4.8). The resulting reflection point is given by $\mathbf{R} = \mathbf{X}(\lambda_0)$. Notice that for non-ruled quadric mirrors equation (4.8) has only one solution.

This method to find the reflection point \mathbf{R} in a noncentral catadioptric vision system presents a major advantage over the method of using explicitly the Euclidean expressions of the mirror either using the Snell's Law (equation (4.3)) or the Fermat Principle (equation (4.6)). This advantage is the fact that, once intersected the quadrics \mathbf{Q} and \mathbf{S} , the solution is given by a nonlinear equation in only one parameter. This is important for the accuracy of the solution and also to the computational efficiency of the method.

4.3 Inverting the projection model

The inverse correspondence between a pixel and a direction in space is also relevant. In this section it is derived the explicit and closed form solution for the spatial line that is projected at a given image point. This expression uses only simple algebra and assumes the knowledge of the mirror quadric and the intrinsic parameters of the camera.

Consider then the camera centered in \mathbf{C} with the intrinsic matrix \mathbf{K} . The quadric is a non-ruled specular surface whose canonical form is given by equation (3.11). If \mathbf{T} represents the generic position and orientation of

the camera in relation to the quadric, composed by a rotation matrix \mathbf{Rot} and a translation vector \mathbf{t} , the quadric mirror is expressed in the camera coordinate system as:

$$\mathbf{Q}_{\text{cam}} = \{q_{ij}\} = \mathbf{T}^{-T} \mathbf{Q} \mathbf{T}^{-1} = \begin{bmatrix} \mathbf{Rot} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix}^{-T} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & A & B/2 \\ 0 & 0 & B/2 & -C \end{bmatrix} \cdot \begin{bmatrix} \mathbf{Rot} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix}^{-1} \quad (4.9)$$

In camera coordinates, the reflection point \mathbf{R} projects to the image plane through the camera center by the projection equation [59]:

$$\mathbf{u} = \frac{1}{\lambda} \mathbf{K} [\mathbf{I} | \mathbf{0}] \mathbf{R} \quad (4.10)$$

Due to ambiguity it is impossible to recover the reflection point from the previous equation but one can invert it to recover the ray reflected by the specular surface, also called emanating ray (see also [94]). This 3×3 vector is parameterized as function of λ by:

$$\mathbf{v} = \lambda \mathbf{K}^{-1} \mathbf{u} \quad (4.11)$$

The reflection point $\mathbf{R}(\lambda) = \begin{bmatrix} \mathbf{v}^T & 1 \end{bmatrix}^T$ is for the value of λ that verifies equation $\mathbf{R}^T \mathbf{Q}_{\text{cam}} \mathbf{R} = 0$. By expanding this expression, it yields the following second degree linear equation in λ :

$$\begin{aligned} & (v_1^2 q_{11} + v_2^2 q_{22} + v_3^2 q_{33} + 2v_2 v_3 q_{23} + 2v_1 v_3 q_{13} + 2v_1 v_2 q_{12}) \lambda^2 + \\ & + (2v_1 q_{14} + 2v_2 q_{24} + 2v_3 q_{34}) \lambda + q_{44} = 0 \end{aligned} \quad (4.12)$$

where here, for convenience, q_{ij} represents the element ij of \mathbf{Q}_{CAM} and v_i represents the i -th component of $\mathbf{K}^{-1} \mathbf{u}$.

The appropriate root of equation (4.12), say λ_0 , is chosen to compute the reflection point $\mathbf{R} = \left[\lambda_0(\mathbf{K}^{-1}\mathbf{u})^T \quad 1 \right]^T$, whose coordinates depend on the image point, the camera parameters and the quadric mirror. The appropriate root to choose depends on the orientation of the vector \mathbf{v} . The point \mathbf{R} parameterized should be the first in the positive direction of the depth coordinate x_3 . So, the root to choose is the one that has smallest absolute value (positive if the vector \mathbf{v} points in the positive direction of the x_3 axis and negative otherwise).

For this particular case, Euclidean coordinates make it easier to compute the incident ray \mathbf{v}_i from the reflected ray \mathbf{v}_r (reverse of equation (4.1)) and the normal vector to the mirror (see section 3.1.4).

The reflected ray is the join of the camera center $\mathbf{C} = \left[0 \quad 0 \quad 0 \quad 1 \right]^T$ and the reflection point \mathbf{R} , oriented from the latter to the former. In Cartesian coordinates it is expressed by the normalized emanating vector, such that:

$$\mathbf{v}_r = -\frac{\lambda_0\mathbf{K}^{-1}\mathbf{u}}{\|\lambda_0\mathbf{K}^{-1}\mathbf{u}\|} \quad (4.13)$$

The normal vector to the quadric at the reflection point is given by the three first coordinates of the tangent plane $\mathbf{\Pi}_N = \mathbf{Q}_{\text{cam}}\mathbf{R}$. The normalized normal is thus given by:

$$\mathbf{n} = \frac{1}{\Delta} \cdot \begin{bmatrix} r_1q_{11} + r_2q_{12} + r_3q_{13} + q_{14} \\ r_1q_{12} + r_2q_{22} + r_3q_{23} + q_{42} \\ r_1q_{13} + r_2q_{23} + r_3q_{33} + q_{34} \end{bmatrix} = \frac{[\mathbf{I}_3|\mathbf{0}] \mathbf{Q}_{\text{cam}}\mathbf{R}}{\|[\mathbf{I}_3|\mathbf{0}] \mathbf{Q}_{\text{cam}}\mathbf{R}\|} \quad (4.14)$$

where Δ is the norm of the vector in the numerator, r_i is the i -th component of the reflection point \mathbf{R} and \mathbf{I}_3 is the 3×3 identity matrix.

Although equation 4.1 express the reflected vector as function of the normal and incident vector, one can compute the incident vector \mathbf{v}_i , by

inverting the equation, as function of the normal and incident vectors. It is then written in the form:

$$\mathbf{v}_i = \mathbf{v}_r - 2(\mathbf{v}_r \cdot \mathbf{n})\mathbf{n} \quad (4.15)$$

As the incident vector and the reflection point are known, the Plücker matrix correspondent to the incident line ℓ_i can then be computed using equation (3.8). The explicit equation depends on the image point coordinates (\mathbf{u}), the camera intrinsic parameters (\mathbf{K}) and the quadric mirror in the camera reference frame (\mathbf{Q}_{cam}).

The incident ray is thus expressed by its Plücker matrix given by:

$$\begin{aligned} \mathbf{L}_i = & \begin{bmatrix} \lambda_0 \mathbf{K}^{-1} \mathbf{u} \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -\frac{\lambda_0 \mathbf{K}^{-1} \mathbf{u}}{\|\lambda_0 \mathbf{K}^{-1} \mathbf{u}\|} - 2 \cdot \left(-\frac{(\lambda_0 \mathbf{K}^{-1} \mathbf{u})^T}{\|\lambda_0 \mathbf{K}^{-1} \mathbf{u}\|} \cdot \frac{[\mathbf{I}_3 | 0] \mathbf{Q}_{\text{cam}} \mathbf{R}}{\|[\mathbf{I}_3 | 0] \mathbf{Q}_{\text{cam}} \mathbf{R}\|} \right) \cdot \frac{[\mathbf{I}_3 | 0] \mathbf{Q}_{\text{cam}} \mathbf{R}}{\|[\mathbf{I}_3 | 0] \mathbf{Q}_{\text{cam}} \mathbf{R}\|} \\ 0 \end{bmatrix}^T - \\ & - \begin{bmatrix} -\frac{\lambda_0 \mathbf{K}^{-1} \mathbf{u}}{\|\lambda_0 \mathbf{K}^{-1} \mathbf{u}\|} - 2 \cdot \left(-\frac{(\lambda_0 \mathbf{K}^{-1} \mathbf{u})^T}{\|\lambda_0 \mathbf{K}^{-1} \mathbf{u}\|} \cdot \frac{[\mathbf{I}_3 | 0] \mathbf{Q}_{\text{cam}} \mathbf{R}}{\|[\mathbf{I}_3 | 0] \mathbf{Q}_{\text{cam}} \mathbf{R}\|} \right) \cdot \frac{[\mathbf{I}_3 | 0] \mathbf{Q}_{\text{cam}} \mathbf{R}}{\|[\mathbf{I}_3 | 0] \mathbf{Q}_{\text{cam}} \mathbf{R}\|} \\ 0 \end{bmatrix} \cdot \begin{bmatrix} \lambda_0 \mathbf{K}^{-1} \mathbf{u} \\ 1 \end{bmatrix}^T \end{aligned} \quad (4.16)$$

where $\mathbf{R} = \begin{bmatrix} \lambda_0 \mathbf{K}^{-1} \mathbf{u} & 1 \end{bmatrix}^T$ and λ_0 is the solution of equation (4.12).

Chapter 5

Method 1: Quadric Mirror Shape Recovery and Calibration of Catadioptric Systems

In this chapter we present a method to recover the shape of the mirror and calibrate the camera parameters and its position and orientation in relation to the mirror. The *a priori* data are the intrinsic parameters of the pinhole perspective camera and a set of point coordinates in the world reference frame (or local for simplicity, without loss of generality). Additionally we address the equations for the use of the apparent mirror contour since it can improve the accuracy of the results. This method is published in the Journal of Optical Engineering [54].

5.1 Problem statement

We consider a catadioptric system made up by a camera (perspective or orthographic) and a quadric surface mirror. The camera is represented by its intrinsic parameters matrix \mathbf{K} (3×3) and the mirror surface by its quadric matrix \mathbf{Q} (4×4). The system is fully noncentral such that the pose of the camera in relation to the mirror is unconstrained and represented by the screw rigid transformation \mathbf{T} (a 4×4 matrix).

\mathbf{Q} is expressed in the quadric coordinate system. It is also expressed in camera coordinate system by \mathbf{Q}_{cam} which is related to the quadric in canonical form by $\mathbf{Q}_{\text{cam}} = \mathbf{T}^T \mathbf{Q} \mathbf{T}$. We also consider the quadric matrix given in its block form:

$$\mathbf{Q}_{\text{cam}} = \begin{bmatrix} \mathbf{Q}_{3\text{cam}} & \mathbf{q}_{\text{cam}} \\ \mathbf{q}_{\text{cam}}^T & q_{44\text{cam}} \end{bmatrix} \quad (5.1)$$

Consider now an object in the scene and a set of 3D points \mathbf{P}_i with known coordinates in the reference frame of the local object (or in the world reference frame). This reference frame is related to the camera coordinate system by the screw rigid transformation \mathbf{H} in such a way that $\mathbf{P}_{i\text{cam}} = \mathbf{H} \mathbf{P}_i$.

The location and pose of the object in relation to the camera (represented by the 4×4 matrix \mathbf{H}), is described by three rotations about the coordinate axes (θ_X , θ_Y and θ_Z) and three translations along the same coordinate axes (t_X , t_Y and t_Z). See figure 5.1.

The goal of this chapter is to describe a method for the estimation of the mirror shape given by the quadric surface \mathbf{Q} , the pose of the camera in relation to the mirror (\mathbf{T}) and the pose of the scene points (expressed in local coordinates) in relation to the camera (\mathbf{H}).

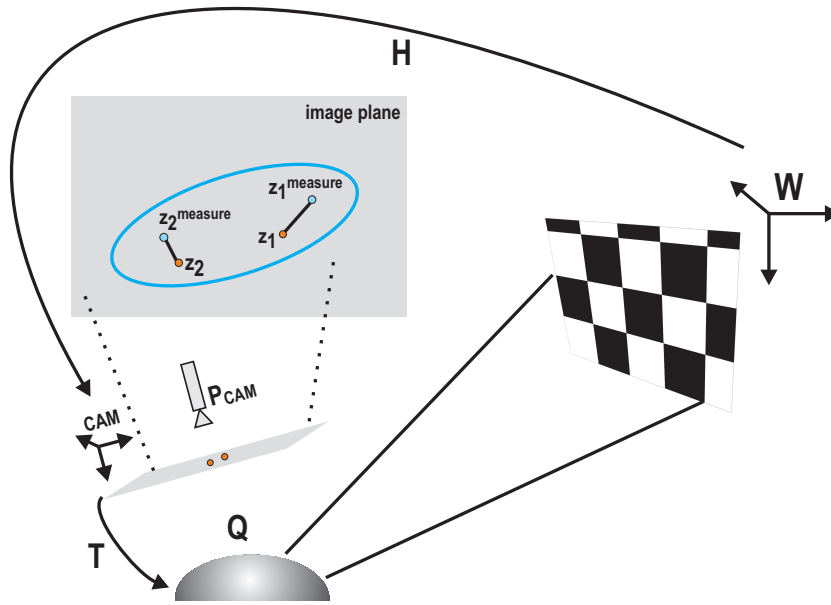


Figure 5.1: Coordinate systems of the camera, world and mirror and their relative positions.

5.2 Apparent Contour

The apparent contour of the quadric mirror contains useful information that can be used to estimate the quadric itself. It reduces the uncertainty in the estimation of the quadric elements.

As stated in [59], under the camera matrix \mathbf{P} the conic \mathbf{C} back-projects to the cone:

$$\mathbf{Q}_{\text{cone}} = \mathbf{P}^T \mathbf{C} \mathbf{P} \quad (5.2)$$

where in the case of a perspective camera we have $\mathbf{P} = \mathbf{K} [\mathbf{I} | \mathbf{0}]$ and in this case it yields:

$$\mathbf{Q}_{\text{cone}} = \begin{bmatrix} \mathbf{K}^T \mathbf{C} \mathbf{K} & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix} \quad (5.3)$$

On the other hand, [59] also states that the cone with vertex \mathbf{V} and tangent to the quadric \mathbf{Q}_{cam} is the degenerate quadric with the equation given by:

$$\mathbf{Q}_{\text{cone}} = (\mathbf{V}^T \mathbf{Q}_{\text{cam}} \mathbf{V}) \mathbf{Q}_{\text{cam}} - (\mathbf{Q}_{\text{cam}} \mathbf{V})(\mathbf{Q}_{\text{cam}} \mathbf{V})^T \quad (5.4)$$

and if the vertex is at the center of the coordinate system ($\mathbf{V} = [0 \ 0 \ 0 \ 1]^T$) and the quadric matrix is expressed in block form, the cone is then given by:

$$\mathbf{Q}_{\text{cone}} = \begin{bmatrix} q_{44\text{cam}} \mathbf{Q}_{3\text{cam}} - \mathbf{q}_{\text{cam}} \mathbf{q}_{\text{cam}}^T & \mathbf{0} \\ \mathbf{0} & 0 \end{bmatrix} \quad (5.5)$$

As a result it can be seen that the cones represented in equations (5.3) and (5.5) are projectively equivalent. That means that the following equations can be written:

$$\left\{ \begin{array}{l} q_{44\text{cam}} q_{11\text{cam}} - q_{14\text{cam}} q_{14\text{cam}} = \Gamma_{11} \\ q_{44\text{cam}} q_{12\text{cam}} - q_{14\text{cam}} q_{24\text{cam}} = \Gamma_{12} \\ q_{44\text{cam}} q_{13\text{cam}} - q_{14\text{cam}} q_{34\text{cam}} = \Gamma_{13} \\ q_{44\text{cam}} q_{22\text{cam}} - q_{24\text{cam}} q_{24\text{cam}} = \Gamma_{22} \\ q_{44\text{cam}} q_{23\text{cam}} - q_{24\text{cam}} q_{34\text{cam}} = \Gamma_{23} \\ q_{44\text{cam}} q_{33\text{cam}} - q_{34\text{cam}} q_{34\text{cam}} = \Gamma_{33} \end{array} \right. \quad (5.6)$$

where Γ_{ij} are the elements of the matrix $\mathbf{\Gamma} = \mathbf{K}^T \mathbf{C} \mathbf{K}$. These six equations correspond to five degrees of freedom since the matrices are symmetric and projectively equivalent. The five degrees of freedom can be represented independently by five ratios. If the last equation is chosen as the reference, the following five equations are obtained:

$$q_{ij\text{cam}} = \frac{\Gamma_{ij}(q_{44\text{cam}}q_{33\text{cam}} - q_{34\text{cam}}q_{43\text{cam}}) + \Gamma_{33}q_{i4\text{cam}}q_{j4\text{cam}}}{q_{44\text{cam}}} \quad (5.7)$$

where $(i, j) \in \{(1, 1), (1, 2), (1, 3), (2, 2), (2, 3)\}$. These equations define five of the quadric mirror parameters as function of the other five, as a function of the intrinsic parameters of the camera and also as a function of the conic that represents the quadric apparent contour.

5.3 Nonlinear calibration of the catadioptric system

This chapter provides an algorithm to calibrate the mirror, its pose in relation to the camera and the pose of the camera in relation to 3D world coordinates. We assume that the intrinsic parameters of the perspective camera are known as well as the local structure of the 3D calibration scene points.

If the apparent contour (which is a conic in the image plane) of the quadric surface mirror is visible (or partially visible) in the image it is possible to estimate its five independent parameters using at least five points (see for instance [59]). This conic is the 3×3 \mathbf{C} matrix in equation (5.3) and since the intrinsic parameters \mathbf{K} are known, the back-projected cone from the apparent contour of the mirror is known. On the other hand, this cone is projectively equivalent to the cone given by equation (5.5), thus providing five equations (5.7), used in our method.

Additionally, as the quadric has nine degrees of freedom instead of the ten which is the number of parameters q_{ij} , the scale factor can be fixed if the value for one parameter is arbitrarily chosen. We may choose for instance $q_{11} = 1$ (however any other value or any other parameter can be chosen) which reduces the number of unknowns and also removes the scale factor

ambiguity. This restriction, however, does not prevent the quadric to be an ellipsoid, a paraboloid or a hyperboloid, the quadrics we study.

The algorithm we propose starts from an initial estimate for all the ten unknown parameters (four for the quadric \mathbf{Q}_{cam} - (q_{14} , q_{24} , q_{34} and q_{44}) and six for the screw \mathbf{H} - three translations and three rotation angles).

Consider the set of image points \mathbf{p}_i and their corresponding 3D coordinates \mathbf{P}_i in the local coordinate system of the calibration object. These 3D points have coordinates in the camera coordinate system given by \mathbf{HP}_i . Back projecting image points to incident lines in space by using geometric optics is trivial, giving a line \mathbf{L}_i for each image point of the set being considered (see section 4.3).

If the quadric \mathbf{Q}_{cam} and the transformation between the camera system and the local object system \mathbf{H} are correct, points \mathbf{HP}_i belong to the lines \mathbf{L}_i . However, in a wrong or noisy configuration, the lines will not pass through the 3D points \mathbf{HP}_i . Thus, the distances between these points and the lines \mathbf{L}_i can be added to define the cost function to be minimized by any nonlinear optimization method.

Once estimated, the three rotation angles and the three translations can be used to compute the transformation matrix \mathbf{H} easily. This step is straightforward.

Although uniqueness of the solution has not been proved, in experiments it is in general achieved.

5.4 Quadric Pose Estimation

In this section we describe how the quadric in its canonical form and its pose in relation to the camera can be estimated starting from the quadric matrix computed in the camera coordinate frame.

The general quadric can be described by a 4×4 symmetric matrix \mathbf{Q}_{cam}

such that $\mathbf{X}^T \mathbf{Q}_{\text{cam}} \mathbf{X} = \mathbf{0}$ holds for all points on the quadric surface. Due to symmetry the quadric has ten coefficients and nine degrees of freedom.

If the quadric parameters q_{11} to q_{44} are computed using some known world structure (as stated by the previous sections) the corresponding quadric matrix can be diagonalized in such a way as to specify a change of coordinates transforming the quadric into its canonical form. That means that the coordinate transformation between the camera and the quadric coordinate frames can be estimated. As a result, the misalignment between the camera and the mirror can also be estimated. In [32, 81] an alternative form to convert the quadric into its canonical form is proposed, using block diagonalization [133, 134]. However this method does not constrain the point transformation to be rigid as required in our case.

Consider a rigid transformation \mathbf{T} made up by a rotation matrix $\mathbf{Rot} = \{r_{ij}\}$ ($i, j \in \{1..3\}$) and a translation vector $\mathbf{t} = \{t_i\}$ ($i \in \{1..3\}$) and the quadric mirror in its reduced form given by equation (5.1). The matrix can be written in the form:

$$\mathbf{T} = \begin{bmatrix} \mathbf{Rot} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \quad \mathbf{Q} = \begin{bmatrix} \mathbf{Q}_3 & \mathbf{q} \\ \mathbf{q}' & q_{44} \end{bmatrix} \quad (5.8)$$

The generic quadric in the camera coordinate system (\mathbf{Q}_{cam}) is obtained through the application of the rigid transformation of the quadric \mathbf{Q} by \mathbf{T} . The relationship is $\mathbf{Q}_{\text{cam}} = \mathbf{T}^T \mathbf{Q} \mathbf{T}$. This equation can be expanded to obtain the following expression:

$$\mathbf{Q}_{\text{cam}} = \begin{bmatrix} \mathbf{Q}_{3\text{cam}} & \mathbf{q}_{\text{cam}} \\ \mathbf{q}_{\text{cam}}^T & q_{44\text{cam}} \end{bmatrix} = \begin{bmatrix} \mathbf{Rot}^T \mathbf{Q}_3 \mathbf{Rot} & \mathbf{Rot}^T \mathbf{Q}_3 \mathbf{t} + \mathbf{Rot}^T \mathbf{q} \\ \mathbf{t}^T \mathbf{Q}_3 \mathbf{Rot} + \mathbf{q}^T \mathbf{Rot} & \mathbf{t}^T \mathbf{Q}_3 \mathbf{t} + 2\mathbf{t}^T \mathbf{q} + q_{44} \end{bmatrix} \quad (5.9)$$

The goal is to estimate linearly the rigid transformation \mathbf{T} and the quadric in its canonical form \mathbf{Q} , starting with the knowledge of the quadric

\mathbf{Q}_{cam} . This estimation is impossible since there are more unknowns (twelve for the transformation and nine for the quadric) than equations. However, some constraints on the quadric \mathbf{Q} allow the recovery of both \mathbf{T} and \mathbf{Q} . The quadric mirror for the most general mirrors can also be expressed in the simpler form:

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & A & B/2 \\ 0 & 0 & B/2 & -C \end{bmatrix} \quad (5.10)$$

and so \mathbf{Q}_3 is a diagonal matrix whose first two diagonal elements are unitary and \mathbf{q} is a 3-vector whose first two components are zero. The mirrors that we want to study are paraboloids, hyperboloids of two sheets, ellipsoids and spheres. The constraints that the parameters must satisfy are: for paraboloids ($C = 0; A = 0$), for ellipsoids ($B = 0$), for hyperboloids ($A < 0; C < 0$) and for spheres ($A = 1; C + B^2/2 > 0$). By changing the coordinates in the Z axis for hyperboloids and spheres, the additional constraint $B = 0$ can be applied. It can be seen that using this framework the parameter B is non zero only for paraboloids and zero for all other mirrors and also that in the former case the parameter A is zero. We thus have:

- paraboloids - $B \neq 0; A = 0; C = 0 \implies \mathbf{Q}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$,

$$\mathbf{q} = \begin{bmatrix} 0 & 0 & B/2 \end{bmatrix}^T \text{ and } q_{44} = 0$$

- hyperboloids, ellipsoids and spheres - $B = 0; A \neq 0 \implies \mathbf{Q}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & A \end{bmatrix}$

$$\text{and } \mathbf{q} = \mathbf{0}^T.$$

The estimation of the rotation matrix \mathbf{Rot} and of the first block diagonal \mathbf{Q}_3 is performed using the diagonalization obtained by the eigenvalue decomposition of the matrix \mathbf{Q}_{3cam} . This step is simple and as a result we obtain two matrices such that $\mathbf{Q}_{3cam}\mathbf{V}_3 = \mathbf{V}_3\mathbf{D}_3$ where \mathbf{D}_3 is a diagonal matrix with the eigenvalues in the principal diagonal and \mathbf{V}_3 is a full matrix with their corresponding eigenvectors as columns. Due to the symmetric nature of quadrics, the eigenvectors matrices are naturally symmetric and therefore it can be written $\mathbf{Q}_{3cam} = \mathbf{V}_3\mathbf{D}_3\mathbf{V}_3^T$ since \mathbf{V}_3 is orthogonal. Furthermore, the diagonal matrix \mathbf{D}_3 can be decomposed in such a way that the central matrix has the elements in the order required in the diagonal as stated in equation (5.10). It thus yields:

$$\mathbf{D}_3 = \begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3 \end{bmatrix} = \mathbf{P}_{ijk} \cdot \begin{bmatrix} d_i & 0 & 0 \\ 0 & d_j & 0 \\ 0 & 0 & d_k \end{bmatrix} \cdot \mathbf{P}_{ijk}^T = \mathbf{P}_{ijk}\mathbf{Q}_3\mathbf{P}_{ijk}^T \quad (5.11)$$

where \mathbf{P}_{ijk} is a permutation matrix possibly necessary to order the eigenvalues in the diagonal. Substituting equation (5.11) into the equation of \mathbf{Q}_{3cam} it then holds that:

$$\mathbf{Rot} = (\mathbf{V}_3\mathbf{P}_{ijk})^T \quad \mathbf{Q}_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & d \end{bmatrix} \quad (5.12)$$

which gives the estimates for the rotation matrix and the first 3×3 block of the quadric. Since the quadric is rotationally symmetric, the first two diagonal elements are equal and can be made, by scaling, equal to 1.

For the remaining unknowns (translation vector \mathbf{t} and q_{44}), the elements \mathbf{q}_{cam} e q_{44cam} are used. However, distinct analysis are made for hyperbolic, elliptic and spherical mirrors and for parabolic mirrors.

For **hyperbolic**, **elliptic** and **spherical** mirrors, in which case $\mathbf{q} = \mathbf{0}^T$, one has:

$$\mathbf{t} = \mathbf{Q}_3^{-1} \mathbf{Rot}^{-T} \mathbf{q}_{cam} \quad (5.13)$$

and

$$q_{44} = q_{44cam} - \mathbf{t}^T \mathbf{Q}_3 \mathbf{t} \quad (5.14)$$

which completes the estimation of the pose of the camera in relation to the quadric mirror for hyperbolic, elliptic and spherical mirrors.

Consider now the case of **parabolic** mirrors in which case $B \neq 0$ and $A = 0; C = 0$. In this case the diagonal matrix \mathbf{Q}_3 has only two non zero elements and so its identification is easy. Expanding now the elements of \mathbf{q}_{cam} the following three linear equations are obtained:

$$\begin{cases} r_{11}t_1 + r_{21}t_2 + r_{31}\frac{B}{2} = q_{1cam} \\ r_{12}t_1 + r_{22}t_2 + r_{32}\frac{B}{2} = q_{2cam} \\ r_{13}t_1 + r_{23}t_2 + r_{33}\frac{B}{2} = q_{3cam} \end{cases} \iff \begin{bmatrix} t_1 \\ t_2 \\ \frac{B}{2} \end{bmatrix} = \mathbf{Rot}^{-T} \mathbf{q}_{cam} \quad (5.15)$$

and finally, expanding the equation for q_{44cam} , the equation to estimate the last unknown is:

$$t_3 = \frac{q_{44cam} - t_1^2 - t_2^2}{B} \quad (5.16)$$

since $q_{44} = 0$.

Therefore we have shown that it is possible to recover the pose of the mirror in relation to the camera and the misalignment of both (as long as the quadric matrix is known in the camera coordinate system). Notice, however, that noise can perturb the identification of parameters whose value is zero. In this case, additional information about the mirror or an exclusion

test can be used to disambiguate (since in our framework $B = 0$ or $A = 0$, exclusively).

5.5 Experiments

We show the usefulness of the method described by applying it to simulated data and also to real images. We estimate the quadric that describes the mirror surface, its pose in relation to the camera and the pose of the camera in relation to the world reference frame.

This section starts by presenting the results with simulated data. Two different camera/mirror configurations are considered: a perspective camera with a spherical mirror with a 37.5mm of radius (the camera and the mirror are not aligned) and an orthographic camera with a parabolic mirror in a paracatadioptric configuration (the axis is aligned with the mirror).

The algorithm tries to minimize the error cost function using two known minimization methods that are combined: a genetic algorithm and the simplex Nelder-Mead method. Successive runs of both methods, one at a time, led to solutions.

Convergence was tested in two different initial conditions. In the first case we added some noise (Gaussian distribution with zero mean) to the ground truth values to use as initial values. The algorithm was then run until one of the stopping conditions was achieved. The stop conditions adopted are:

1. The error is smaller than a tolerance value (e.g. $1e^{-5}$);
2. A predefined maximum elapsed time was reached;
3. Unchanged error value for a predefined time interval (local minimum).

In the second case we start the experiment by using a random vector for the initial values of the parameters to be estimated. We performed several

runs of the algorithm, usually between 10 and 100, using different random starting values and, in the general case, when the stopping conditions are met the estimates obtained are close to the optimal values. Next, these results are used to rerun the genetic algorithm as starting chromosomes. The algorithm was again iterated until one of the stopping conditions was met.

We noticed no difference between the results in both tests. This indicates that the convergence is in general obtained irrespective of using or not values close to the ground truth.

Table 5.1 presents median values of the results obtained for the parameters and the corresponding ground truth values (perspective camera with spherical mirror). Table 5.2 presents the same analysis for the paracatadioptric system.

As it can be seen from the results with simulated data, the quadric in the camera coordinates \mathbf{Q}_{cam} (which integrates both the quadric \mathbf{Q} in the canonical form and the pose \mathbf{T}) and the pose \mathbf{H} are estimated accurately. The estimates for the actual values of \mathbf{Q} and \mathbf{T} are not computed for simulated data since the equations presented are in closed form and their values are highly accurate. They are however computed in the experiments with real images.

In the experiments with real images, we present the results obtained with a noncentral catadioptric system made up by a perspective camera and a spherical mirror. The same camera is also used in a hyperbolic configuration. The camera used was a commercial CANON EOS 350D with image size 3456x2304. The spherical mirror has a radius of 37.5mm. The systems were previously calibrated using a two step algorithm by first estimating the perspective camera parameters and then calibrating the mirror. The mirrors were not perfectly aligned with the camera. Figure 5.2 displays an image acquired by both cameras.

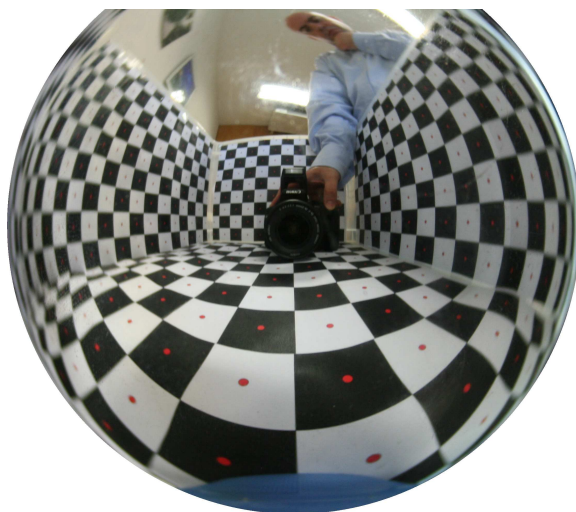
The calibration object used to calibrate the camera and to test the al-

Table 5.1: Tests with simulated data for a perspective camera with an off-axis spherical mirror configuration. This table presents median values of the results. Notice that the parameters of the error random distribution (top row) are in relation to the ground truth values of the parameters. GT means ground truth, $N(\cdot)$ means normal error distribution and $Unif(\cdot)$ means uniform error distribution.

	GT	GT+err($\sim N(0; 0.5)$)	Random $\sim Unif(0; 1)$	Random $\sim Unif(0; 2)$	Random $\sim Unif(-1; 1)$	Random $\sim Unif(-2; 2)$
q_{14}	32.0	33.18	31.78	32.23	33.23	31.06
q_{24}	16.0	16.83	15.28	17.31	17.36	17.09
q_{33}	-150.0	-147.14	-149.95	-152.48	-153.04	-144.10
q_{34}	2.000e4	2.190e4	2.247e4	2.256e4	2.146e4	2.301e4
θ_X	0.0	-0.01	-0.01	0.01	0.00	-0.01
θ_Y	0.0	0.00	0.01	0.01	0.00	0.01
θ_Z	0.0	0.00	0.01	0.01	-0.01	0.00
t_X	0.00	0.00	-0.00	-0.01	-0.01	-0.01
t_Y	0.0	0.01	0.00	-0.00	0.01	-0.02
t_Z	-300.0	-296.84	-303.14	-298.59	-298.82	-299.47

Table 5.2: Tests with simulated data. The configuration is made up of an orthographic camera with an aligned parabolic mirror. This table presents median values of the results. Notice that the parameters of the error random distribution (top row) are in relation to the ground truth values of the parameters. GT means ground truth, $N(\cdot)$ means normal error distribution and $Unif(\cdot)$ means uniform error distribution.

	GT	GT+err($\sim N(0; 0.5)$)	Random $\sim Unif(0; 1)$	Random $\sim Unif(0; 2)$	Random $\sim Unif(-1; 1)$	Random $\sim Unif(-2; 2)$
q_{14}	0.0	0.00	-0.00	0.00	0.01	0.01
q_{24}	0.0	0.01	-0.00	-0.01	-0.02	0.00
q_{33}	-600.0	-587.39	-605.24	-600.39	-599.15	-589.31
q_{34}	1.800e5	1.785e5	1.781e5	1.780e5	1.778e5	1.843e5
θ_X	0.0	-0.00	-0.00	0.01	-0.01	0.00
θ_Y	0.0	-0.01	-0.00	0.01	-0.01	0.01
θ_Z	0.0	-0.01	-0.00	-0.00	0.00	0.01
t_X	0.0	0.01	0.02	-0.02	-0.00	-0.01
t_Y	0.0	0.00	-0.01	0.00	0.00	-0.01
t_Z	-300.0	-296.52	-292.72	-308.60	-298.01	-299.61



(a) Spherical mirror



(b) Hyperbolic mirror

Figure 5.2: Real images used.

gorithm can be planar or non-planar. Non-planar calibration objects usually yield better calibration results since their non-planar structure inherently provides additional information. However, tests performed with planar patches proved to provide estimates with good accuracy. Planar patterns have the advantage of simplicity. We chose to use planar patterns in the experiments with the hyperbolic mirror and with the spherical mirror we used non-planar objects to test the method.

Additionally, since in the case of the hyperbolic mirror the apparent contour of the quadric mirror is not visible (because the mirror is a hyperboloid cut by a plane) we present the results without using the apparent contour information.

First we present the results concerning the spherical mirror configuration. As in the case of simulated data, the accuracy of the method is tested in two ways namely by using initial values for the parameters that are obtained by adding some noise to the true values and also by starting the iterations with random values for the parameters. Monte Carlo methods are applied and the results obtained are presented in table 5.3 for the case of the spherical mirror using the apparent contour and in table 5.4 without using the apparent contour.

Using the hyperbolic mirror, the experiments are similar and their results are presented in table 5.5.

The results obtained in the experiments with real images show that estimates for the quadric mirror parameters in the camera coordinate system (\mathbf{Q}_{cam}) and for the pose of the camera in the world reference frame (\mathbf{H}) can be obtained with good accuracy. The use of the apparent contour allows for a drastic improvement of the accuracy of the estimates.

The next step is the estimation of the parameters of the quadric mirror in its canonical form (\mathbf{Q}) and the estimation of its pose in relation to the camera - (\mathbf{T}). For that purpose we use the equations of section 5.4. Since the mirror

Table 5.3: Tests with real images obtained from a perspective camera with a **spherical** mirror. Median values of estimated parameters for the quadric mirror in the camera coordinate system - \mathbf{Q}_{cam} and for the pose of the camera in relation to the world reference frame - \mathbf{H} , using the **apparent contour** to reduce the uncertainty. Notice that the parameters of the error random distribution (top row) are in relation to the ground truth values of the parameters. GT means ground truth, $N(,)$ means normal error distribution and $\text{Unif}(,)$ means uniform error distribution.

	GT	GT+err($\sim N(0; 0.5)$)	Random $\sim \text{Unif}(0; 1)$	Random $\sim \text{Unif}(0; 2)$	Random $\sim \text{Unif}(-1; 1)$	Random $\sim \text{Unif}(-2; 2)$
q_{14}	-0.18	0.01	0.56	0.64	0.17	0.16
q_{24}	11.67	11.75	9.97	10.41	11.30	11.04
q_{33}	-272.46	-269.54	-269.30	-280.18	-286.35	-279.64
q_{34}	7.296e4	7.140e4	7.128e4	7.716e4	8.056e4	7.686e4
θ_X	0.00	-0.00	0.00	0.00	-0.00	-0.00
θ_Y	0.00	-0.00	-0.00	-0.00	-0.00	-0.00
θ_Z	0.00	0.00	0.00	0.00	0.00	0.00
t_X	50.00	50.71	51.90	51.97	50.80	50.83
t_Y	50.00	50.41	49.47	49.08	49.77	50.04
t_Z	-300.00	-302.60	-303.18	-293.31	-287.40	-293.53

Table 5.4: Tests with real images obtained from a perspective camera with a **spherical** mirror. Median values of the estimated parameters for the quadric mirror in the camera coordinate system - \mathbf{Q}_{cam} and for the pose of the camera in relation to the world reference frame - \mathbf{H} , not using the **apparent contour**. Notice that when the apparent contour is not used the algorithm only converges to the solution with good initial parameter estimates. Also notice that the parameters of the error random distribution (top row) are in relation to the ground truth values of the parameters. GT means ground truth, $N(\cdot)$ means normal error distribution and $\text{Unif}(\cdot)$ means uniform error distribution.

	GT	GT+err~ $N(0; 0.5)$	Random~ $\text{Unif}(0; 1)$	Random~ $\text{Unif}(0; 2)$	Random~ $\text{Unif}(-1; 1)$	Random~ $\text{Unif}(-2; 2)$
q_{11}	1.00	1.00	-4.18	-51.86	0.97	-11.59
q_{12}	0.00	-0.00	0.36	-0.41	-0.00	0.37
q_{13}	0.00	0.00	1.21	-1.34	0.00	1.75
q_{14}	-0.18	-0.19	0.72	-37.40	-0.28	229.81
q_{22}	1.00	1.00	-4.63	-53.23	0.97	-12.10
q_{23}	0.00	0.00	1.13	-1.92	0.00	2.16
q_{24}	11.67	11.50	0.92	44.17	7.74	404.01
q_{34}	-272.46	-268.90	0.11	-268.98	-187.28	-161.46
q_{44}	72964.16	71074.54	-0.30	-22139.85	34502.16	-66674.35
θ_X	0.00	0.00	-0.08	3.02	0.00	-0.08
θ_Y	0.00	0.00	0.21	-3.03	0.00	0.08
θ_Z	0.00	0.00	0.03	-3.13	0.00	0.01
t_X	50.00	49.99	-34.41	108.32	49.74	20.44
t_Y	50.00	50.12	107.27	122.11	52.95	111.54
t_Z	-300.00	-303.25	-505.44	-551.32	-378.04	-663.26

Table 5.5: Tests with real images obtained from a perspective camera with a **hyperbolic** mirror. Median values of the estimated parameters for the quadric mirror in the camera coordinate system - \mathbf{Q}_{cam} and for the pose of the camera in relation to the world reference frame - \mathbf{H} , without using the **apparent contour**. Notice that the parameters of the error random distribution (left column) are in relation to the ground truth values of the parameters. GT means ground truth, $N(\cdot)$ means normal errors distribution and $\text{Unif}(\cdot)$ means uniform error distribution.

	GT	GT+err \sim $N(0; 0.005)$	GT+err \sim $N(0; 0.01)$	GT+err \sim $N(0; 0.05)$	GT+err \sim $N(0; 0.1)$	GT+err \sim $N(0; 0.5)$	GT+err \sim $N(0; 1.0)$
q_{11}	1.00	1.00	1.01	0.93	1.59	2.60	-10.06
q_{12}	0.00	0.00	0.00	0.01	-0.03	0.15	-8.24
q_{13}	0.00	-0.00	-0.00	0.03	0.07	-0.01	1.16
q_{14}	4.11	4.45	4.48	-1.74	-11.67	11.30	-2323.02
q_{22}	0.00	-0.00	0.00	-0.01	-0.06	-0.01	0.29
q_{23}	-5.14	-5.14	-5.18	-3.43	17.92	-7.77	-1480.27
q_{24}	-0.76	-0.76	-0.75	-0.73	-0.42	-1.12	-0.06
q_{34}	138.02	137.82	137.32	143.04	80.84	194.75	166.29
q_{44}	-24436.7	-24358.40	-24356.89	-27326.14	-10872.63	-29178.30	-357550.1
θ_X	0.00	0.00	-0.01	-0.01	-0.35	0.11	-0.29
θ_Y	0.00	0.00	0.01	-0.02	0.07	0.08	-0.05
θ_Z	0.00	-0.00	0.00	-0.02	0.09	0.02	0.26
t_X	50.00	47.95	43.60	56.52	47.74	4.09	-166.45
t_Y	50.00	47.92	54.09	53.88	244.74	33.07	254.14
t_Z	-300.00	-300.96	-299.00	-277.94	-152.44	-307.41	-177.92

Table 5.6: Experimental results with real images obtained by a perspective camera with a **spherical** mirror. Estimated values for the pose of the mirror in relation to the camera - \mathbf{t} and for its radius, using the **apparent contour**.

	t_1	t_2	t_3	<i>Radius</i>	Rel. err (%)
GT	-0.1837	11.667	-272.46	37.5	•
GT+err($\sim N(0; 0.5)$)	0.01	11.64	-266.87	37.14	0.97
Random ($\sim Unif(0; 1)$)	0.56	9.92	-267.73	37.01	1.30
Random ($\sim Unif(0; 2)$)	0.64	10.34	-278.45	38.51	2.70
Random ($\sim Unif(-1; 1)$)	0.16	11.22	-284.22	39.45	5.20
Random ($\sim Unif(-2; 2)$)	0.16	10.96	-277.57	38.53	2.74

is a sphere, the translation between the camera and the center of the mirror and its radius are the ones that are meaningful, since the rotation parameters do not change the sphere position and orientation for vision system purposes. Tables 5.6 and 5.7 show the results obtained for the calibration parameters when the apparent contour is used and not used respectively. The rotation matrix \mathbf{Rot} is hence not shown since due to the nature of the mirror any other rotation would express the same mirror (as the orientation of a sphere is meaningless).

From the results one can conclude that the radius of the mirror surface is obtained with high accuracy in all cases where the apparent contour is used. In the cases where the apparent contour is not used, accurate estimates for the radius can only be obtained if the initial values for the parameters are close to their optimal values.

For the hyperbolic mirror, the results for the estimates of the quadric mirror parameters and for its pose in relation to the camera are presented in table 5.8.

It should be remarked that the results in experiments showed that a good

Table 5.7: Experimental results with real images obtained by a perspective camera with a **spherical** mirror. Estimated values for the pose of the mirror in relation to the camera - \mathbf{t} and for the radius, without using the **apparent contour**. Notice that when the apparent contour is not used the algorithm only converges to the solution if started with good initial estimates.

	t_1	t_2	t_3	<i>Radius</i>	Rel. err (%)
GT	-0.1837	11.667	-272.46	37.5	•
GT+err($\sim N(0; 0.5)$)	-0.19	11.50	-268.90	36.98	1.40
Random ($\sim Unif(0; 1)$)	5.09	9.42	-7.79	18.54	50.55
Random ($\sim Unif(0; 2)$)	-1.07	79.39	-50.32	58.03	54.76
Random ($\sim Unif(-1; 1)$)	-0.28	7.74	-187.27	25.19	32.83
Random ($\sim Unif(-2; 2)$)	-19.32	13.76	262.19	39.78	6.09

Table 5.8: Experimental results with real images obtained by a perspective camera with a **hyperbolic** mirror. Estimated values for the pose of the mirror in relation to the camera - \mathbf{t} and for the radius, without using the **apparent contour**. Notice that when the Gaussian error introduced is very high, the algorithm does not converge to an useful solution.

	θ_1	θ_2	θ_3	t_1	t_2	t_3
GT	0.0	0.0	0.0	4.11	-5.12	-181.42
GT+err($\sim N(0; 0.005)$)	0.00	-0.00	-0.40	4.87	-4.49	-182.12
GT+err($\sim N(0; 0.01)$)	0.00	-0.00	0.48	0.83	-7.30	-183.33
GT+err($\sim N(0; 0.05)$)	-0.00	0.18	-0.31	13.83	0.75	-190.05
GT+err($\sim N(0; 0.1)$)	0.01	-0.01	0.12	2.73	-9.33	-203.45
GT+err($\sim N(0; 0.5)$)	0.77	-1.40	0.88	2.02	54.50	0.44
GT+err($\sim N(0; 1.0)$)	0.51	-0.37	0.02	21.99	179.46	-3159.41

estimate for the conic corresponding to the apparent contour of the mirror is very important for the accuracy of the results. This is the most sensitive set of parameters and small errors in their values can cause much higher errors in the estimation of all the other parameters.

Reprojection error

In order to have an exact idea of how erroneous is the estimation and what implications it does have in the calibrated model, we reproject the 3D points to the image and compare the actual and recomputed pixel locations. Figure 5.3 shows some cases of reprojected points and the actual pixels, using some results obtained with the spherical mirror, with and without the apparent contour information.

Consider the state vector as the estimated parameters vector composed by the pose elements of the camera in relation to the mirror and in relation to the world reference frame and also by the mirror shape parameters.

Table 5.9 presents some additional cases and the average errors per pixel. Table 5.9 presents results obtained with all the spherical mirror data (contour and no contour) and with the hyperbolic mirror. The corresponding amplitude and angle error of the estimated state vector are presented instead of the relative errors for each parameter, as presented in the previous tables, for data compactness.

As seen in table 5.9, the reprojection error per pixel grows with the error in the estimated values. However, it can also be observed that even to poor accurate state vectors (such that the amplitude and angle errors are above 5% – 10%), the reprojection error is sometimes low. This can be explained by the fact that the parameters contribute heterogeneously for the image formation and these estimated state vectors conduce to a reasonably good projection. At this stage it is, however, difficult to infer the differences in the influence of the calibration parameters concerning the projection.

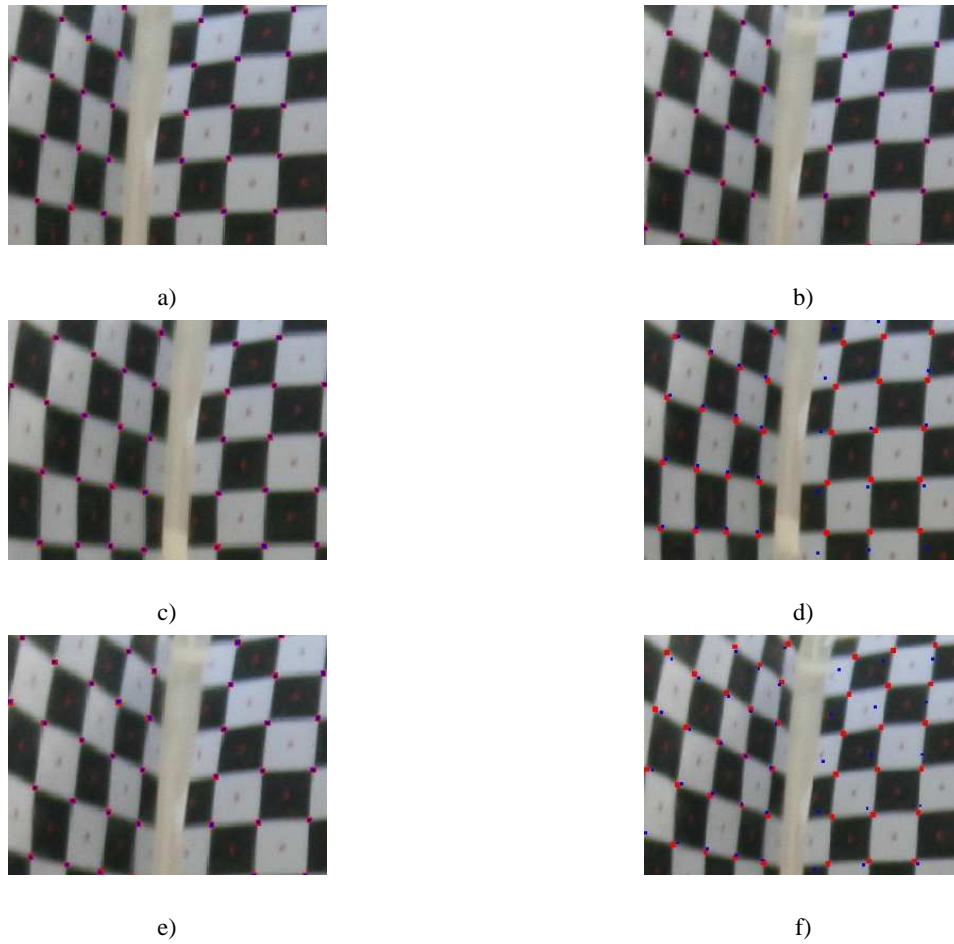


Figure 5.3: Reprojection of the 3D points in the image. The left column presents results obtained using the apparent contour and the right column the results obtained without using the contour. In (a)-(b) the initial estimate is the ground truth vector with Gaussian noise added, with standard deviation of 50% of the actual value, in (c)-(d) the initial estimate is a random vector between 0 and 1 (0 \rightarrow 100% of the actual value) and in (e)-(f) the initial estimate is a random vector between -1 and 1 (-100% \rightarrow 100% of the actual value).

Table 5.9: Reprojection error of 3D points (in pixels) and the corresponding amplitude error (in percentage) and angle errors (in degrees) of the estimated state vectors.

Mirror	Contour	Initial estimate	Amplitude error	Angle error	Reproj. error
Spherical	Yes	GT+N(0;0.5)	0.15	0.16	0.01
Spherical	No	GT+N(0;0.5)	1.34	1.46	0.03
Spherical	Yes	Random $\in [0, 1]$	0.28	0.32	0.02
Spherical	No	Random $\in [0, 1]$	525.43	92.62	61.18
Spherical	Yes	Random $\in [0, 2]$	1.83	2.16	0.16
Spherical	No	Random $\in [0, 2]$	626.60	74.87	62.81
Hyperbolic	No	GT+N(0;0.005)	0.36	15.48	0.02
Hyperbolic	No	GT+N(0;0.01)	3.99	6.66	0.04
Hyperbolic	No	GT+N(0;0.05)	27.14	37.55	0.13
Hyperbolic	No	GT+N(0;0.1)	119.12	71.61	0.38
Hyperbolic	No	GT+N(0;0.5)	502.50	89.74	9.98
Hyperbolic	No	GT+N(0;1.0)	385.62	66.98	12.00

5.6 Summary and Conclusions

In this chapter we presented a method to calibrate catadioptric systems made up by a perspective or orthographic camera (whose internal parameters are considered to be known) and a curved mirror whose shape is mathematically expressed by a non-degenerate quadric (includes spherical, hyperbolic, parabolic and elliptic mirrors). The method requires the knowledge of the intrinsic parameters of the camera and also local world calibration information (for instance distances in a calibration pattern).

The method allows for the use of the apparent contour of the mirror (which is a conic) to constrain the quadric mirror and its pose in relation to the camera and then applies a nonlinear iterative minimization method to match some back projected pattern points to a 3D grid. This method first estimates the pose of the camera in relation to the world reference frame and the quadric mirror in camera coordinates. In the second step, using closed form expressions, it estimates the camera in its canonical form and its pose.

Experimental results showed that the method is accurate both with simulated data and with real images, even when the initial estimates (required by the nonlinear optimization procedure) are completely random, specially if the apparent contour is used. The calibration objects used in the experiments can be planar and non-planar. It was concluded that the estimation of the conic parameters corresponding to the mirror apparent contour is critical to the accuracy of the results.

To understand how the estimation errors in the parameters affects the calibration of the catadioptric system, we reprojected the 3D points to the image. It was observed that when the apparent contour was used (and if their conic parameters was acquired with accuracy), the reprojection error is small, indicating a good calibration result. However, if the contour was not used, the results are more degraded.

In what concerns the computer efficiency of the method, for the minimization of the cost function, we used two nonlinear minimization algorithms simultaneously: the simplex Nelder-Mead method and a genetic algorithm. Both methods use a strategy based on several iterations providing a slow convergence and so they are computationally demanding. Besides that, they provide well-behaved convergence.

Finally, our conclusion about this method is that it provides trustable results if some conditions are met. These conditions are the accurate calibration of the primary optical element, the camera, by its intrinsic parameters \mathbf{K} , extremely accurate computation of the conic elements of the apparent contour \mathbf{C} and non-planar calibration objects are used.

Chapter 6

Method 2: Estimating Parameters of Noncentral Catadioptric Systems using Bundle Adjustment

In this chapter we present a new method for the estimation of the parameters of a noncentral catadioptric system using bundle adjustment techniques. The key idea is to relax Snell's Law to a set of incident light rays, projecting their intersecting point with the quadric mirror to the image and minimizing the reprojection error as done in the usual bundle adjustment method for camera calibration. The relaxation of Snell's Law is necessary since there is no closed-form projection model for these vision systems (see chapter 4).

We are interested in the estimation of the intrinsic and extrinsic parameters of general catadioptric systems with quadric mirrors regardless of being central or not. The method is composed by two steps. In the first one the system is calibrated in the sense of a black box model, that is, we assume

that the correspondence pixel \longleftrightarrow 3D line is provided by using Grossberg and Nayar method [55], Sturm and Ramalingam method [124] or by some other method. We opted to use known motion between dense calibration grids to perform a stable ray calibration.

The second calibration step proposed by the method is the application of the class of bundle adjustment methods for camera calibration to general (central or not) catadioptric vision systems. The explicit computation of the Jacobian of the projection equations is possible due to the relaxation of Snell's law constraint. The non existence of closed-form equations for the projection (and hence the non existence of a way to provide an estimate for the coordinates of the reflection point on the mirror surface) is circumvented by the fact that there are available correspondences between pixels and lines in space and not between pixels and points in space. The intersections between the direction rays and the mirror surface thus provide the reflection points.

Bundle adjustment is then applied to the projection model by using the following parameterization: intrinsic parameters of the pinhole camera (5 parameters), position and orientation of the camera in the world reference frame (three rotation angles and three displacements - 6 parameters), the quadric mirror shape parameters in canonical form (3 parameters) and the position and orientation of the camera in relation to the mirror (three rotation angles and three displacements - 6 parameters). The total number of parameters of the state vector is 20. We show that bundle adjustment methods are suitable for the calibration of general catadioptric systems and that the convergence is generally achieved both in experiments with simulated data and in experiments with real images. Since bundle adjustment methods require an initial estimate for the state vector, we also provide an automatic algorithm to compute the initial estimates.

Rotations are parameterized by Euler angles. As is widely known, usually

Euler angles present stability and numerical problems due to their high non linear nature. This problem is solved by using frozen (or cumulative) and update rotation matrices in the bundle adjustment optimization algorithms. The current estimate in each iteration is frozen in a rotation matrix and the derivatives are evaluated in the update angles rather than in the accumulated ones. This strategy provides very simple Jacobian expressions. After the update state vector is computed, the next iteration starts by accumulating the last update in the frozen angles.

This method was published in the 6th Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras Workshop in conjunction with the International Conference on Computer Vision (October 2005, Beijing, China) [52]. An extended and revised version with several enhancements was submitted to the Computer Vision and Image Understanding Journal (CVIU) and is now under revision.

6.1 Problem statement

Consider the camera coordinate system as the reference frame. Consider now a catadioptric vision system made up of a pinhole camera whose intrinsic parameters are given by matrix \mathbf{K} :

$$\mathbf{K} = \begin{bmatrix} f_u & \nu & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (6.1)$$

and a specular mirror surface given by the quadric in its canonical form:

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & A_s & B_s \\ 0 & 0 & B_s & -C_s \end{bmatrix} \quad (6.2)$$

The camera is positioned in the center of the main reference frame and its poses (position and orientation) in relation to the quadric mirror and to the world reference frame are given by \mathbf{T} and \mathbf{H} respectively.

$$\mathbf{T} = \begin{bmatrix} \mathbf{Rot}(\theta_C, \theta_B, \theta_A) & \mathbf{t}_T \\ \mathbf{0}^T & 1 \end{bmatrix} \quad (6.3)$$

and

$$\mathbf{H} = \begin{bmatrix} \mathbf{Rot}(\theta_3, \theta_2, \theta_1) & \mathbf{t}_H \\ \mathbf{0}^T & 1 \end{bmatrix} \quad (6.4)$$

where θ_A , θ_B and θ_C are the rotation Euler angles of the pose camera-mirror and θ_1 , θ_2 and θ_3 are the rotation Euler angles of the pose camera-world. The rotation matrices \mathbf{Rot} are the concatenation of elementary rotation matrices around a single axis. $\mathbf{t}_T = \{t_{Ti}\}$ and $\mathbf{t}_H = \{t_{Hi}\}$ with $i \in \{1, 2, 3\}$ are the translational vectors for the poses considered.

Consider now a set of correspondences between pixel (u, v) in the image plane and a 3D line expressed in the world coordinate system. The lines in space are represented by two points \mathbf{A} and \mathbf{B} also expressed in the world coordinate frame. Figure 6.1 shows the relative position of the camera and the set of lines in space representing the incident light rays.

The aim of this method is to achieve the calibration of the intrinsic and extrinsic parameters of the general catadioptric vision system. Using the framework presented so far, the parameters to calibrate are then \mathbf{K} , \mathbf{Q} , \mathbf{T} and \mathbf{H}

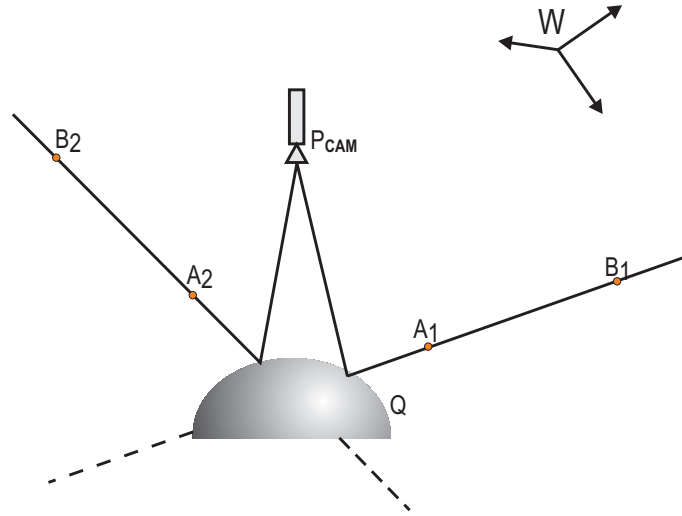


Figure 6.1: World reference frame and the catadioptric system.

Bundle adjustment methods require one cost function that depend on the parameters to be estimated. The computation of its Jacobian either by analytical expressions or numerical derivations can be used to enhance the accuracy and performance of the methods. Usually this class of methods are used with projection models that map 3D points in space into image points. The cost function is often the weighted sum of the squared errors which are the Euclidean difference between the measured position of the point in the image and the estimated or predicted one. However, in our framework we do not use points to project but rather use correspondences between lines in 3D-space and points in the image plane. Furthermore, additional information can be used to enhance the convergence of the method. Other methods that do not employ Jacobian can also be used. In the next section we derive the equations for the projection model 3D line \longleftrightarrow image point and its Jacobian.

6.2 Projection model

An explicit projection equation for a general catadioptric vision system does not exist, although for the particular case of central cameras this expression is known. Such lack is due to the nature of specular reflection and image formation, which are highly nonlinear, well modeled by an implicit equation.

Specular reflection is modeled by Snell's Law. According to Snell's Law the incident and reflection angles are equal. If we backproject a pixel, the reflection ray is easily found out (by means of an explicit closed-form equation).

However, if an arbitrary incident ray in the specular surface is considered, it may or may not be projected through the optical center of the camera and imaged at a pixel. Additionally, when projecting the ray by using the reflection laws, there are usually multiple mathematical solutions and the selection of the real one is sometimes not trivial.

In our model this Snell's Law constraint is relaxed to allow the analytic computation of a corresponding image pixel to any ray in space that intersects the specular surface. Without this simplification and considering the projection model to be the real specular case, most of the incident rays would not be imaged (since they would not pass at the camera optical center) and then there would not be an error measurement to minimize.

Our projection model is rather simplified by considering the projection in the image of the reflection point, computed as the intersection between the 3D line in space and the quadric surface as shown in figure 6.2. The error resulting from not considering Snell's Law is taken into account by incorporating it in the bundle adjustment cost function. This is done by computing the difference between the unit directions of the reflected rays according to Snell's Law and to our projection model (by directly projecting it into the image relaxing the projection law).

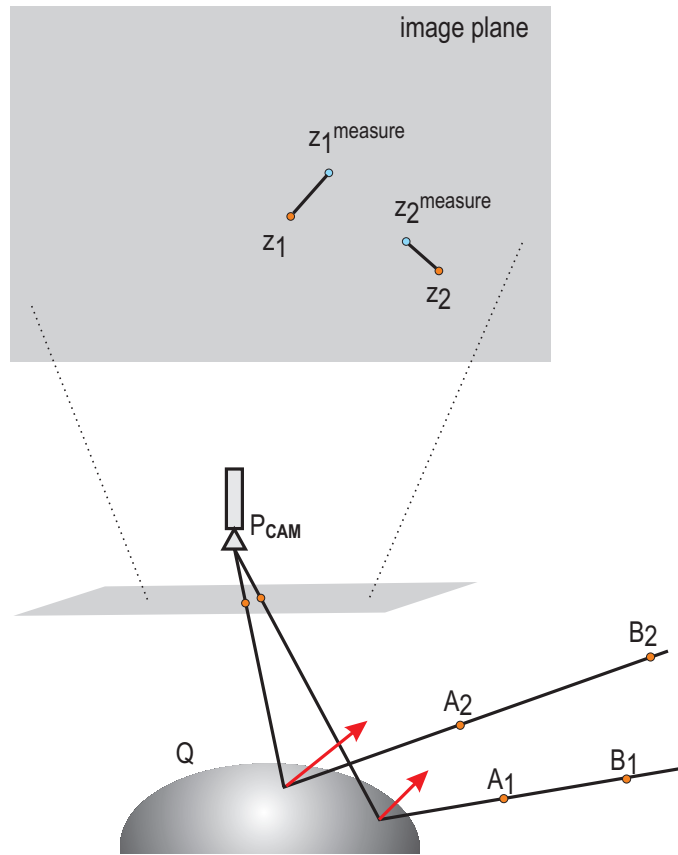


Figure 6.2: Projection model mapping lines in space (given by couples of points A_i and B_i) to pixels in the image. Notice that Snell's law is relaxed so that the incident and reflection angles are not constrained to be equal.

Before computing the reflection point \mathbf{R}_{CAM} and projecting it into image, we must first express every geometrical entity in the camera reference system. Points $\mathbf{A}_{\mathbf{w}}$ and $\mathbf{B}_{\mathbf{w}}$ (3×1 vectors) belonging to the incident ray are expressed in world coordinates. Their expressions in the camera coor-

dinate system are simply $\begin{bmatrix} \mathbf{A} & 1 \end{bmatrix}^T = \mathbf{H} \begin{bmatrix} \mathbf{A}_w & 1 \end{bmatrix}^T$ and $\begin{bmatrix} \mathbf{B} & 1 \end{bmatrix}^T = \mathbf{H} \begin{bmatrix} \mathbf{B}_w & 1 \end{bmatrix}^T$.

The mirror surface, expressed in the canonical form by \mathbf{Q} , is also expressed in the camera coordinate system as $\mathbf{Q}_{\text{CAM}} = \mathbf{T}^{-T} \mathbf{Q} \mathbf{T}^{-1}$ (notice that in relation to the method in previous chapter we opted to use here the transformation \mathbf{T} in the inverse sense). By replacing matrices \mathbf{T} and \mathbf{Q} by their expressions, one obtains bilinear equations that depend on the quadric shape parameters (A_s , B_s and C_s) and on the camera-mirror pose transformation (t_{T1} , t_{T2} , t_{T3} , θ_A , θ_B and θ_C). The explicit expressions are presented in appendix.

The reflecting point \mathbf{R}_{CAM} is then a point belonging to the line that join points \mathbf{A} and \mathbf{B} and also that is on the quadric surface \mathbf{Q}_{CAM} . Its expression can be written in the form $\mathbf{R}_{\text{CAM}} = \mathbf{A} + \alpha(\mathbf{B} - \mathbf{A})$ where the parameter α is the solution of the equation corresponding to incidence on the quadric:

$$\begin{aligned} \begin{bmatrix} \mathbf{R}_{\text{CAM}}^T & 1 \end{bmatrix} \mathbf{Q}_{\text{CAM}} \begin{bmatrix} \mathbf{R}_{\text{CAM}} \\ 1 \end{bmatrix} &= \\ &= \begin{bmatrix} \mathbf{A}^T + \alpha(\mathbf{B} - \mathbf{A})^T & 1 \end{bmatrix} \mathbf{Q} \begin{bmatrix} \mathbf{A} + \alpha(\mathbf{B} - \mathbf{A}) \\ 1 \end{bmatrix} = 0 \end{aligned} \quad (6.5)$$

Equation (6.5) is a quadratic equation on α that depends on the quadric mirror coefficients and also on points \mathbf{A} and \mathbf{B} that define the line in space. The reflection point on the mirror surface is then obtained by selecting the appropriate root of the quadratic equation.

The selection of the appropriate root is also important to the projection model and depends on the relative position of points \mathbf{A} and \mathbf{B} and on their position in relation to the quadric surface too. Generally the approach is to admit some simple relationship over the points and quadric that permits

us to choose the right root. As the points \mathbf{A} and \mathbf{B} are computed from the incident ray and they are usually points in scene objects, we admit that they are both actually imaged (that is to say that they are not in the occluded or virtual part of the ray). If an additional assumption is made, that the point \mathbf{A} is closer to the quadric than the point \mathbf{B} (which is easily achieved by interchanging them if this is not true), then the appropriate root is simply the one with the smallest absolute value. Another approach is to project both points and choose the one with the smallest reprojection error.

Once obtained the reflection point \mathbf{R}_{CAM} , its projection in the image plane is given by:

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \lambda \mathbf{K} \mathbf{R}_{CAM} \quad (6.6)$$

where \mathbf{K} is the matrix of the intrinsic parameters and λ is the scale factor. Since we are interested in the image coordinates themselves, to eliminate the scale factor we divide the first two coordinates by the third. Expanding their equations it yields:

$$\begin{cases} u = \frac{z_1}{z_3} = \frac{f_u R_{CAM_1} + \nu R_{CAM_2} + u_0 R_{CAM_3}}{R_{CAM_3}} \\ v = \frac{z_2}{z_3} = \frac{f_v R_{CAM_2} + v_0 R_{CAM_3}}{R_{CAM_3}} \end{cases} \quad (6.7)$$

which is the projection model for an arbitrary catadioptric system considering that the correspondences between pixels \longleftrightarrow incident lines are provided. We emphasize that this projection model is an approximation since it relaxes Snell's Law of reflection. This is done in order to obtain closed form projection equations to use in bundle adjustment (a complete closed form expression is achieved once selected the appropriate root of a second-order polynomial). The errors due to the approximation in this model are minimized by the nonlinear optimization algorithm, as discussed in section 6.4.

6.3 Ray Calibration

The first step of the algorithm is performed by calibrating the rays that are subsequently used in the state vector estimation.

As mentioned in the introduction, the ray calibration is made by applying a known motion transformation to the camera (or to the scene) and by estimating for each pixel the new world coordinates of the point imaged. This is done by inverting a homography between 3D and image planes as described in [124]. We briefly review it.

Consider a pixel with image coordinates (u, v) . Consider that in the first image the 3D point that is projected into image has known world coordinates \mathbf{A} . Suppose now that a given motion transformation \mathbf{Mot} is applied to the camera (or to the objects in the scene as a whole) such that the coordinates of the points became $\hat{\mathbf{A}} = \mathbf{Mot} \cdot \mathbf{A}$ and that the n neighboring points in the second image are $\hat{\mathbf{B}}_i$ with $i \in \{1..n\}$ with also known world coordinates ($\hat{\mathbf{B}}_i = \mathbf{Mot} \cdot \mathbf{B}_i$). Consider that the image points corresponding to the n neighboring points are (u_i, v_i) .

To calibrate the line that is projected by the vision system to the pixel (u, v) one has to know at least two points of the line. One of the points is already known, that is, the point \mathbf{A} and the other one can be calculated by interpolating the coordinates of the points in its neighborhood \mathbf{X}_{3D} .

This calculation assumes that the projection is continuous in a neighborhood of the points, say a radius of rad pixels, and that all points are coplanar (although the image is actually from a curved surface). The homography \mathbf{Hom} relates coplanar points as $\mathbf{x}_{img} = \mathbf{Hom}\mathbf{X}_{3D}$ or:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} hom_{11} & hom_{12} & hom_{13} \\ hom_{21} & hom_{22} & hom_{23} \\ hom_{31} & hom_{32} & hom_{33} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad (6.8)$$

The homography is calculated by expanding this equation for the n neighboring pixels and estimating its nine parameters by using:

$$\begin{bmatrix} X_i & Y_i & Z_i & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & X_i & Y_i & Z_i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & X_i & Y_i & Z_i \\ & & & \vdots & & & & & \end{bmatrix} \begin{bmatrix} hom_{11} \\ hom_{12} \\ hom_{13} \\ \vdots \\ hom_{32} \\ hom_{33} \end{bmatrix} = \begin{bmatrix} u_i \\ v_i \\ 1 \\ \vdots \end{bmatrix} \quad (6.9)$$

with $i \in \{1..n\}$. With three (non collinear) or more points the homography is uniquely estimated and the coplanar 3D points that corresponds to the image pixel $\mathbf{x}_{\text{img}} = [u \ v \ 1]^T$ is easily estimated by $\mathbf{X}_{3D} = \mathbf{Hom}^{-1} \mathbf{x}_{\text{img}}$.

Once obtained these two 3D points for every pixel (the point in the scene before the motion - \mathbf{A}_i and the scene estimated point after the motion - $\hat{\mathbf{B}}_i$) it is straightforward to compute the corresponding rays since they are simply the join of those two points.

The number of points used to estimate the homography is important to the accuracy of this ray calibration, as well as how far they are from the given pixel. This distance of the pixels on image is itself dependent on the type and amplitude of motion. In the section of experiments we discuss the accuracy of this ray calibration.

6.4 Bundle adjustment

Bundle adjustment methods are generally suitable for large scale problems with a large number of variables and often with a high degree of non-linearity. In general a nonlinear iterative multidimensional minimization algorithm is applied to the state vector starting from an initial position. The function to

be minimized, the cost function, is usually a sum of squared errors between predicted and measured positions in the image plane. There are several minimization strategies (see [132] for a detailed discussion) based on the derivatives of the cost function - the Jacobian, since the problem is multidimensional.

In our problem, the cost function is the sum of squared errors given by:

$$f(\mathbf{x}) = \sum_1^N \{W_I \Delta \mathbf{z}(\mathbf{x})^T \Delta \mathbf{z}(\mathbf{x}) + W_A \Delta \mathbf{V}_r(\mathbf{x})^T \Delta \mathbf{V}_r(\mathbf{x})\} \quad (6.10)$$

where $\Delta \mathbf{z}(\mathbf{x}) = \mathbf{z}(\mathbf{x}) - \mathbf{z}^{\text{measure}}$ and $\mathbf{z}(\mathbf{x}) = \begin{bmatrix} u & v \end{bmatrix}^T$ (computed by equation (6.7)) and $\Delta \mathbf{V}_r(\mathbf{x})$ is the difference between the unit reflected rays computed by the projection model of section 6.2 and computed according to Snell's Law. W_I and W_A are weight values applied to the reprojection error (W_I) and to the incident ray direction error (W_A). \mathbf{x} represents the state vector whose elements are as follows:

$$\mathbf{x} = [f_u \ f_v \ \nu \ u_0 \ v_0 \ t_{H1} \ t_{H2} \ t_{H3} \ \theta_1 \ \theta_2 \dots \\ \dots \theta_3 \ A_s \ B_s \ C_s \ t_{T1} \ t_{T2} \ t_{T3} \ \theta_A \ \theta_B \ \theta_C] \quad (6.11)$$

Consider that $\mathbf{V}_r = -\mathbf{R}_{\text{CAM}} / \|\mathbf{R}_{\text{CAM}}\|$ is the unit reflected ray computed according to the model presented in section 6.2 and that \mathbf{V}'_r is the actual reflected ray computed according to Snell's Law. Assume that they are expressed in the camera coordinate frame. Their expressions are:

$$\left\{ \begin{array}{l} \mathbf{V}_r = -\frac{1}{\sqrt{(A_1 + \alpha(B_1 - A_1))^2 + (A_2 + \alpha(B_2 - A_2))^2 + (A_3 + \alpha(B_3 - A_3))^2}} \begin{bmatrix} A_1 + \alpha(B_1 - A_1) \\ A_2 + \alpha(B_2 - A_2) \\ A_3 + \alpha(B_3 - A_3) \end{bmatrix} \\ \mathbf{V}'_r = \mathbf{V}_i - 2(\mathbf{V}_i^T \mathbf{N}) \mathbf{N} \end{array} \right. \quad (6.12)$$

where \mathbf{V}_i is the unit incident ray and \mathbf{N} is the normal vector to the quadric surface at the reflection point \mathbf{R} . For the incident ray we have:

$$\mathbf{V}_i = \frac{\mathbf{A} - \mathbf{B}}{\|\mathbf{A} - \mathbf{B}\|} \quad (6.13)$$

To compute the normal vector to the quadric we take into account that the normal to the quadric is the direction vector of the tangent plane at the reflection point \mathbf{R}_{CAM} . Hence, the tangent plane is given by $\mathbf{\Pi}_R = \mathbf{Q}_{\text{CAM}} \begin{bmatrix} \mathbf{R}_{\text{CAM}}^T & 1 \end{bmatrix}^T$ and since the direction vector of the plane $\mathbf{\Pi}_R$ is made up by the first three components of it, we have $\mathbf{N} = \mathbf{Q}_{\text{CAM}_{3 \times 4}} \begin{bmatrix} \mathbf{R}_{\text{CAM}}^T & 1 \end{bmatrix}^T$, where $\mathbf{Q}_{\text{CAM}_{3 \times 4}}$ is the rectangular matrix made up by the first three lines of the quadric mirror matrix \mathbf{Q}_{CAM} .

The second element of the cost function can then be computed using the following equation:

$$\Delta \mathbf{V}_r(\mathbf{x}) = \mathbf{V}'_r - \mathbf{V}_r \quad (6.14)$$

The computation of the Jacobian is straightforward by taking the derivatives of the equations of the cost function with respect to each of the unknowns of the state vector. The derivative for each term is computed independently and then summed as $\mathbf{J} = \mathbf{J}_I + \mathbf{J}_A$. The explicit expressions for the Jacobian are presented in appendix for the sake of clarity. They are however straightforward to derive.

For the error component of the cost function corresponding to the image reprojection - $W_I \Delta \mathbf{z}(\mathbf{x})^T \Delta \mathbf{z}(\mathbf{x})$, its expanded expression for each point used is given by $W_I \left((u(\mathbf{x}) - u^{\text{measure}})^2 + (v(\mathbf{x}) - v^{\text{measure}})^2 \right)$. Taking the derivatives of this equation with respect to the components of the state vector \mathbf{x} , it yields the first component of the Jacobian, expressed as:

$$\mathbf{J}_I = 2W_I \begin{bmatrix} (u - u^{measure}) \frac{\partial u}{\partial f_u} + (v - v^{measure}) \frac{\partial v}{\partial f_u} \\ (u - u^{measure}) \frac{\partial u}{\partial f_v} + (v - v^{measure}) \frac{\partial v}{\partial f_v} \\ (u - u^{measure}) \frac{\partial u}{\partial v} + (v - v^{measure}) \frac{\partial v}{\partial v} \\ (u - u^{measure}) \frac{\partial u}{\partial u_0} + (v - v^{measure}) \frac{\partial v}{\partial u_0} \\ (u - u^{measure}) \frac{\partial u}{\partial v_0} + (v - v^{measure}) \frac{\partial v}{\partial v_0} \\ (u - u^{measure}) \frac{\partial u}{\partial t_{H1}} + (v - v^{measure}) \frac{\partial v}{\partial t_{H1}} \\ \dots \\ (u - u^{measure}) \frac{\partial u}{\partial \theta_C} + (v - v^{measure}) \frac{\partial v}{\partial \theta_C} \end{bmatrix} \quad (6.15)$$

The Jacobian of the error due to the deviation from Snell's Law \mathbf{J}_A is also calculated by taking the derivatives of the expression of the $\Delta \mathbf{V}_r$ with respect to the components of the state vector, yielding:

$$\mathbf{J}_A = 2W_A \begin{bmatrix} (V_{r1} - V'_{r1}) \frac{\partial V_{r1} - V'_{r1}}{\partial f_u} + (V_{r2} - V'_{r2}) \frac{\partial V_{r2} - V'_{r2}}{\partial f_u} + (V_{r3} - V'_{r3}) \frac{\partial V_{r3} - V'_{r3}}{\partial f_u} \\ (V_{r1} - V'_{r1}) \frac{\partial V_{r1} - V'_{r1}}{\partial f_v} + (V_{r2} - V'_{r2}) \frac{\partial V_{r2} - V'_{r2}}{\partial f_v} + (V_{r3} - V'_{r3}) \frac{\partial V_{r3} - V'_{r3}}{\partial f_v} \\ \dots \\ (V_{r1} - V'_{r1}) \frac{\partial V_{r1} - V'_{r1}}{\partial \theta_C} + (V_{r2} - V'_{r2}) \frac{\partial V_{r2} - V'_{r2}}{\partial \theta_C} + (V_{r3} - V'_{r3}) \frac{\partial V_{r3} - V'_{r3}}{\partial \theta_C} \end{bmatrix} \quad (6.16)$$

Several optimization methods exist that, using the Jacobian, iterate in the state vector space until convergence to a minimum. The most used are the Newton and the Levenberg-Marquardt methods. While the former is easier to implement, the latter is more suitable when numerical instabilities perturb the solution and also when matrix $\mathbf{J}^T \mathbf{J}$ is singular or near singular. As described in the experimental section, we chose to use the Levenberg-Marquardt method to minimize the cost function since we generally obtained better results.

Additional information can be used either to restrict the problem by reducing the number of parameters or to help the convergence by introducing some third term in the cost function. As presented in section 5.2 the apparent contour of the whole mirror can provide this additional data, since

it introduces some algebraic constraints in the quadric mirror parameters. Alternative equations, both for the cost function and for the Jacobian are derived and presented. Another possible enhancement to the convergence is the introduction of a regularization term in the cost function.

6.4.1 Initial Estimate

Regarding the initial estimate for the minimization method, usually some information about the camera and the mirror is provided by the manufacturer. However, no information is in this case available for the position and orientation of the camera in the world reference frame. Although the information available can enhance the quality and precision of the first estimate, we wish to evaluate the robustness of the algorithm without this kind of data.

A totally automatic algorithm to provide the first estimate has obvious advantages. We will next present some possible options to compute one initial estimate and we emphasize that this initial estimate can be enhanced whenever additional information is available.

Initial Estimate Algorithm

Step 1 - Mirror/optical axis. The optical axis is coincident with the mirror axis in a rotationally symmetric system. If we assume a rotationally symmetric vision system, to estimate the optical axis we then must provide two points on the mirror axis.

Before computing the mirror/optical axis and two points on it, consider the point closest to all the incident lines in space, say \mathbf{M}'_0 . In a central catadioptric system this point is the actual viewpoint. However, as our system is noncentral, this point is not the viewpoint.

It can also be observed that the optical axis which is coincident with the mirror axis (since we are assuming a rotationally symmetric system for the

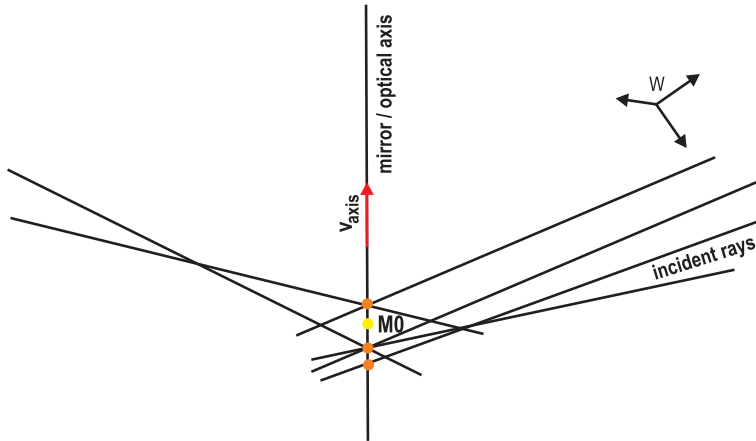


Figure 6.3: Algorithm for the initial estimate - Mirror/optical axis.

computation of the initial estimate) is the line that intersects all the incident lines, that is, all incident rays pass through the optical axis.

Picking up four random lines, Hohmeyer and Teller [63] presented an easy way to compute the line that passes through these four lines. This line should be the mirror/optical axis. An easy adaptation to the method can be made to compute the line that passes closer to all the incident lines (in the least squares sense) or instead of that, one can pick some sets of random lines (amongst all incident rays) and then compute the line that minimizes the square distances. This line should also pass through M'_0 but, in practice, this doesn't happen due to approximation and noise. The closest point to the mirror/optical axis is then computed, say M_0 . Figure 6.3 illustrates the computation of the mirror/optical axis. Consider the unitary direction of the mirror/optical axis given by v_{axis} .

Step 2 - Mirror type, dimension and positioning. Although the type of the mirror is sometimes known we assume that we do not have this information. Since in practice the difference of the type of mirror in the

image formation is high we propose to try different mirror configurations and at the end choose the best one in terms of output value of the cost function. Formally, quadric mirrors are parameterized by three parameters: A_s , B_s and C_s . Generally, for their canonical representation, one may say that for a paraboloid B_s is the only parameter different from zero (positive or negative), for a sphere $A_s = 1$ and $C_s > 0$, for a general ellipsoid $B_s = 0$ and for a hyperboloid $A_s < 0$ and $C_s < 0$.

For simplicity we propose to consider $B_s = 0$ and iterate through some reasonable intervals for A_s and C_s , either considering the mirror a hyperboloid or a sphere. Notice that C_s is the square of the radius for a sphere and it is also the square radius of the circular section for a hyperboloid when the clipping plane is perpendicular to the mirror/optical axis). Once determined these intervals all subsequent steps of this algorithm are computed once for each (A_s, C_s) pair.

As for the position of the mirror center along the mirror/optical axis, for hyperboloids we consider the projection as being central and then all incident rays pass through the focus of the mirror. We consider this point to be the previously computed \mathbf{M}_0 . The center of the mirror is computed by adding the focal distance ($focal = \sqrt{C_s/A_s - C_s}$) to the focus \mathbf{M}_0 in the \mathbf{v}_{axis} direction: $\mathbf{C}_{hyp} = \mathbf{M}_0 + \sqrt{C_s/A_s - C_s}\mathbf{v}_{axis}$. Figure 6.4 illustrates this construction.

For spherical mirror, on the other hand, the parameters $A_s = 1$ but there is no focal point to be placed in \mathbf{M}_0 . We propose to add an additional parameters, say d , to place the center of the sphere so that $\mathbf{C}_{sph} = \mathbf{M}_0 + d\mathbf{v}_{axis}$. The scalar d must be smaller than the radius of the mirror (in absolute value) and thus can be chosen in the interval $d \in [-\sqrt{C_s}, \sqrt{C_s}]$. The iteration of the subsequent steps of the algorithm are performed to each pair of (d, C_s) . Figure 6.5 illustrates this construction.

At the end of this step the quadric mirror in canonical form is known -

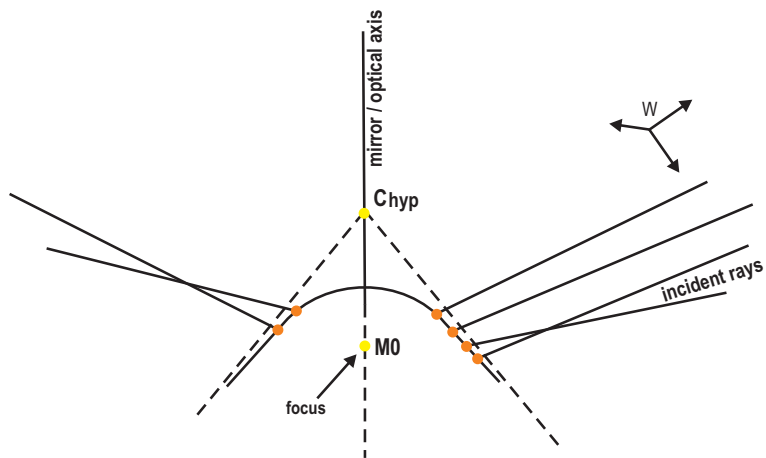


Figure 6.4: Algorithm for the initial estimate - Center of the hyperbolic mirror.

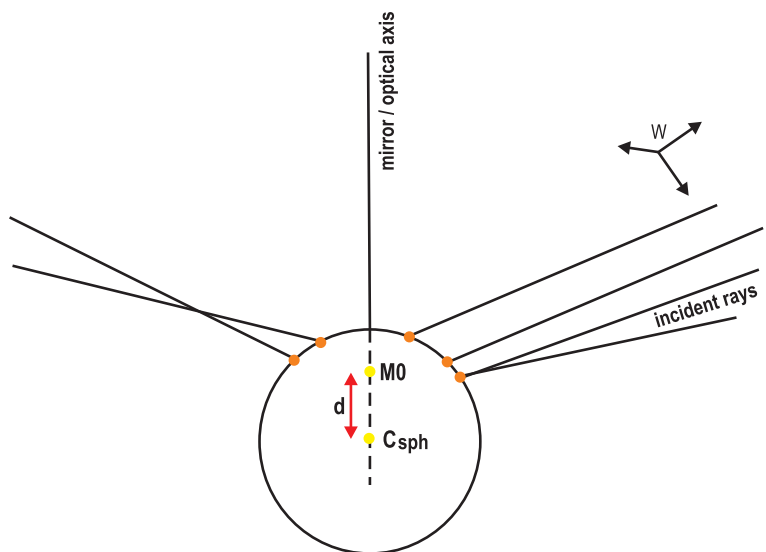


Figure 6.5: Algorithm for the initial estimate - Center of the spherical mirror.

\mathbf{Q} , as well as the center of the mirror in world coordinates.

Step 3 - Quadric mirror in world coordinates. After having picked up two values for A_s and C_s for the hyperbolic case or for d and C_s for the spherical one, we can now compute the transformation between the quadric and world reference frame, and compute the quadric matrix in world frame too.

For the hyperbolic case, the \mathbf{v}_{axis} direction should be transformed to be coincident with the z-axis so we compute two angles (θ_X and θ_Y) and perform a compound rotation around X and Y axis respectively, say \mathbf{R}_q . The translation vector \mathbf{t}_q can be computed by $\mathbf{t}_q = -\mathbf{R}_q \mathbf{C}_{\text{hyp}}$ and the transformation between the world reference frame and the quadric mirror one is then given by:

$$\mathbf{H}_q = \begin{bmatrix} \mathbf{R}_q & \mathbf{t}_q \\ \mathbf{0}^T & 1 \end{bmatrix} \quad (6.17)$$

where the quadric mirror in its canonical form is given by:

$$\mathbf{Q}_{\text{hyp}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & A_s & 0 \\ 0 & 0 & 0 & C_s \end{bmatrix} \quad (6.18)$$

and consequently the quadric mirror is expressed in world coordinates as $\mathbf{Q}_w = \mathbf{H}_q^T \mathbf{Q}_{\text{hyp}} \mathbf{H}_q$.

For the spherical case, the transformation matrix between the mirror and world reference frames has only translation components, since it is meaningless to rotate a sphere. The translation vector is the vector between the center of the sphere and the world origin, so we have $\mathbf{t}_q = -\mathbf{C}_{\text{sph}}$. The transformation is then:

$$\mathbf{H}_q = \begin{bmatrix} \mathbf{I}_3 & \mathbf{C}_{\text{sph}} \\ \mathbf{0}^T & 1 \end{bmatrix} \quad (6.19)$$

where \mathbf{I}_3 is the 3×3 identity matrix and where the quadric mirror in its canonical form is given by:

$$\mathbf{Q}_{\text{sph}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & C_s \end{bmatrix} \quad (6.20)$$

and then the spherical mirror in the world coordinate system is also expressed as $\mathbf{Q}_w = \mathbf{H}_q^T \mathbf{Q}_{\text{sph}} \mathbf{H}_q$.

Step 4 - Reflection points. The computation of the reflection point on the mirror surface is easily computed by intersecting the incident rays with the quadric \mathbf{Q}_w . It can be achieved by computing the values of the parameter α that solve expression:

$$\left[\mathbf{A}_w^T + \alpha (\mathbf{B}_w - \mathbf{A}_w)^T \quad 1 \right] \mathbf{Q}_w \left[\mathbf{A}_w^T + \alpha (\mathbf{B}_w - \mathbf{A}_w)^T \quad 1 \right]^T = 0$$

Generally two solutions exist and as discussed in section 6.2 the appropriate is the smallest one in absolute value.

This step provides a set of points on the mirror surface expressed in world coordinates.

Step 5 - Perspective camera calibration. The position and orientation of the camera in relation to the world reference frame (extrinsic parameters) and its intrinsic parameters can then be estimated by using the known world coordinates of the reflection points on the mirror surface. There are several methods to perform this calibration step. We propose to use the

linear approach described in [38] (p. 45) and the possible refinements also described.

This perspective camera calibration gives us all the camera-world pose parameters (t_{H1} , t_{H2} , t_{H3} , θ_1 , θ_2 and θ_3) and also the intrinsic parameters of the camera (f_u , f_v , ν , u_0 and v_0). Initial values for matrices \mathbf{H} and \mathbf{K} are at this point estimated.

Step 6 - Camera-mirror pose. The position and orientation of the camera in relation to the mirror (matrix \mathbf{T}) can then be estimated from the two poses \mathbf{H}_q and \mathbf{H} and the quadric in world coordinates, since:

$$\begin{aligned}\mathbf{Q}_{\text{cam}} &= \mathbf{T}^{-T} \mathbf{Q} \mathbf{T}^{-1} \\ &= \mathbf{H}^{-T} \mathbf{Q}_w \mathbf{H}^{-1} = \mathbf{H}^{-T} \mathbf{H}_q^T \mathbf{Q} \mathbf{H}_q \mathbf{H}^{-1}\end{aligned}\quad (6.21)$$

and therefore $\mathbf{T} = \mathbf{H} \mathbf{H}_q^{-1}$. The quadric mirror in camera coordinates can then be computed by expression (6.21).

Step 7 - Computation of the cost function. Since a complete state vector is at this point computed, it is used to compute the cost function value (reprojection error and Snell's Law deviation). This value is saved.

Step 8 - Iteration. Iterates steps 3 to 7 for all pairs of parameters (A_s, C_s) for the hyperbolic mirror hypothesis and (d, C_s) for the spherical one. The pairs are picked up from the intervals defined in step 2.

Step 9 - Selection. Select the initial estimate as the one amongst all state vectors tried with the smallest cost.

End of algorithm.

All the parameters of the state vector are now computed and an initial estimate exists. This automatic algorithm to provide an initial set of estimates for the bundle adjustment method can be improved if additional information about the system is available.

6.4.2 Apparent Contour

As stated by section 5.2, from the apparent contour of the quadric mirror useful data for the algorithm can be extracted. It allows the reduction of the uncertainty in the estimation of the parameters of the quadric.

The equations derived enable us to compare the real apparent contour with the analytic contour resulting for the state vector in use. Alternatively we can compare the back-projected and projection cones of equations (5.3) and (5.5) which yields more compact equations since the matrix of the intrinsic parameters \mathbf{K} does not need to be inverted. These equations can be used in the model by introducing a third term in the cost function. The extended explicit equations are presented in the appendix B (section 6.6).

6.5 Experiments

In this section we present the results from the experiments performed to test the robustness and accuracy of the framework presented throughout the chapter. Results with simulated data are first presented. They focus on the robustness of the convergence and on the accuracy of the initial estimates of the state vector. Finally we present some results from experiments with real images using two noncentral catadioptric systems.

6.5.1 Experiments with Simulated Data

In the experiments with simulated data we used a catadioptric systems made up of a pinhole camera with a hyperbolic mirror. Tests with spherical and

parabolic mirrors were also performed but since the results are similar to those obtained with the hyperbolic mirror we omit their results.

We present tests for the two steps of the method: the ray calibration and the state vector estimation. Finally we also present some results obtained by using the method as a whole.

For the first step of the method, that is, the ray calibration, we generated a second image of two planar grid patches and computed the incident rays as explained in section 6.3. For the computation of the homography we used the four closest points in the neighborhood of the image pixel considered ($n = 4$). To evaluate the accuracy of this method to estimate lines corresponding to image pixels we computed the angle between the estimated line and the ground truth one. As the accuracy is affected mainly by the density of grid pixels, expressed by the average image distance to the considered pixel, figure 6.6 plots the RMS value of the angle as function of the average distance of the 4 used points to compute homography.

As can be observed from figure 6.6, as the density of grid pixels decreases, the error in estimating the incident rays that correspond to image pixels increases inversely. This shows that the accuracy of the ray calibration is affected by the assumption of continuous projection of a planar surface. This fact also suggests that when calibrating the incident ray space in this way one should provide high density of pixels in image such that the homography calculated is as close as possible to the actual transformation.

For the calibration of the state vector, the algorithm second step, we evaluated its accuracy by adding Gaussian white noise to the data in two different sets of tests. In the first one we added noise to the ground truth values of the parameters \mathbf{x} . This test evaluates the robustness of the convergence near the optimal point. The energy of the error added to the ground truth values of the state vector was increased from a low error energy (standard deviation of 0.125% of the ground truth value) and was successively

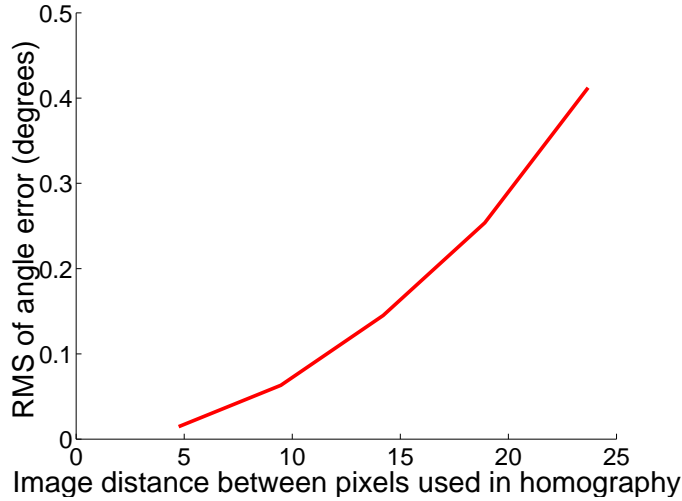


Figure 6.6: Angle error (RMS value in degrees) between the computed and actual incident rays in 3D space. The angle error is computed for an increasing value of the RMS distance between the 4 image pixels used in the computation of the homography.

multiplied by a factor of 2 until a high error energy (standard deviation of 4% of the ground truth value). Figures 6.7 and 6.8 shows the relative error for each of the parameters of the state vector. The test was repeated 100 times and RMS values are presented. For the parameters whose ground truth value is zero we omitted their estimated values since no relative error can be computed. These parameters are ν , θ_2 and B_s . We observed however, that their estimated values were around zero and that their absolute value drifted from zero as the error energy increases, as expected.

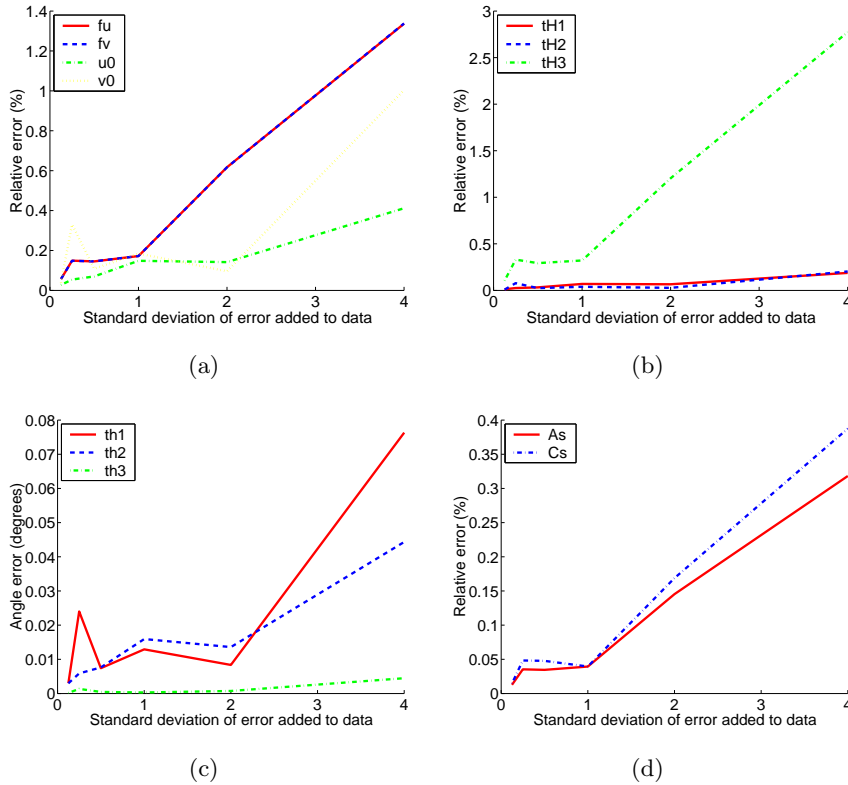


Figure 6.7: Results with **simulated data**. Relative error, or angular error for Euler angles, of the estimated state vector elements, for increasing noise energy added to the initial estimate. The results are plotted in percentage - % and degrees for angles. The parameters whose ground truth value is zero are omitted since no relative error can be computed - ν , θ_2 and B_s . The standard deviation of the noise added is expressed in percentage of the ground truth value.

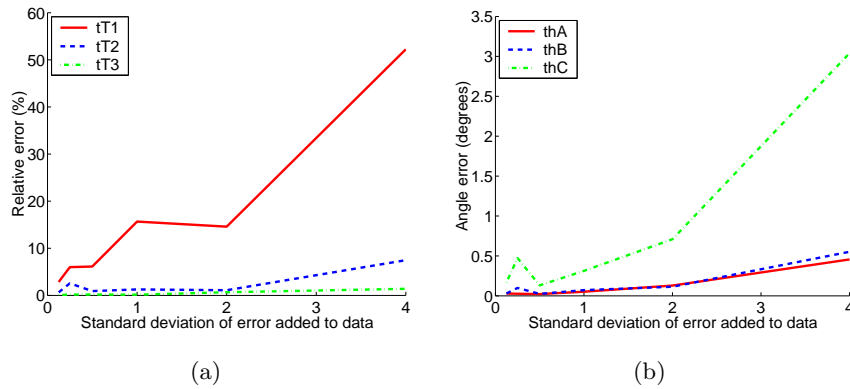


Figure 6.8: **Cont.** - Results with **simulated data**. Relative error, or angular error for Euler angles, of the estimated state vector elements, for increasing noise energy added to the initial estimate. The results are plotted in percentage - % and degrees for angles. The parameters whose ground truth value is zero are omitted since no relative error can be computed - ν , θ_2 and B_s . The standard deviation of the noise added is expressed in percentage of the ground truth value.

The aim of this test is the study of the behavior of the algorithm near the global minimum of the state vector. As can be observed from the figure, the convergence to the ground truth value (or at least a value with very low cost) is generally obtained. To evaluate also the importance of the errors of the state vector parameters in the image itself, we plotted in figure 6.9 the RMS value of the reprojection error and Snell's Law deviation of incident lines. To compute these error measures we did not use the points over which we iterated the bundle adjustment but rather used the remaining point of the calibration patches. As expected, their values increase with noise energy, however, very slowly.

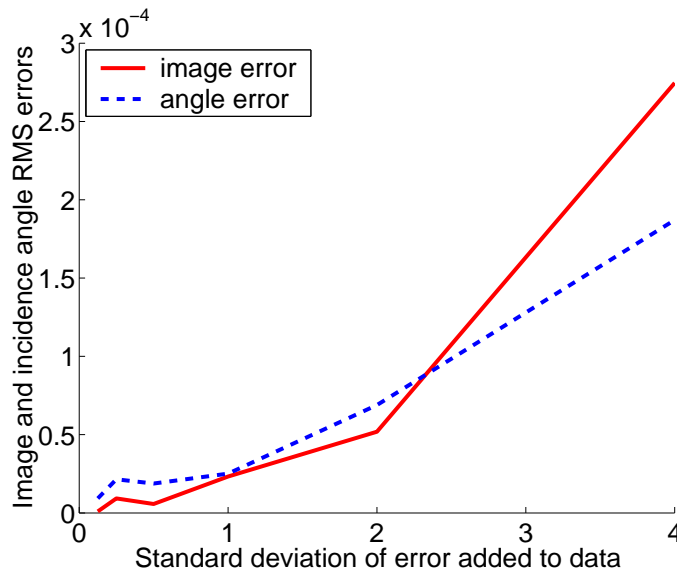


Figure 6.9: Results with **simulated data**. Image error and angle deviation from Snell’s Law obtained with the estimated state vectors. These error measures were computed in points not used in the bundle adjustment process. The standard deviation of the noise added is expressed in percentage of the ground truth value.

The second test was performed by starting to iterate the algorithm from a totally random position. The values were picked up randomly from a wide interval. The test interval is a range from -200% to 200% of the ground truth value (eg. for a parameter whose value is 10 we consider a range from -10 to 30). In this case, as expected, often the starting estimate is a non physical configuration (such that not all or even none of the points are projected into the image). We rejected these cases and picked up another starting estimate. Also, often the random initial values do not converge to the ground truth state vector, getting trapped in a local minimum, corresponding to a high value of the cost function. These cases were also rejected. We concluded

that when totally random state vector estimates are used, the convergence is difficult and slow. However, if the test is repeated a high number of times the chance of obtaining a cost function with low value and therefore convergence is reasonable (on average one has to try about 200 random state vectors to achieve convergence). No results are shown in this case, since when the convergence is achieved, the relative errors of the parameters are small, regardless of the number of trials required to achieve it.

Table 6.1: Standard deviations used to add noise to the image coordinates (in pixels) and to the state vector parameters (in percentage of the ground truth value) in the experiments of the whole method with simulated data.

Image coordinates	[pixel]	0.0625	0.125	0.25	0.5	1.0
State vector par.	%	0.625	1.25	2.5	5	10

As for the experiments with the method as a whole, including the evaluation of the accuracy obtained by the initial estimate method, the two steps of the algorithm were tested in conjunction. Pixels of corresponding calibration patches were picked up from two images taken by the vision system in two different positions affected by a known motion transformation. The image coordinates of the pixel at both images were affected by a noise of increasing standard deviation and the incident lines in space were then estimated. The error was a zero mean Gaussian noise with minimum standard deviation of 0.0625 pixels and maximum of 1 pixel (see table 6.1). These incident rays were used to calibrate the state vector using the second step of the algorithm, the bundle adjustment. Two different evaluations were performed. In the first one the state vector starts from a random position in the neighborhood of their ground truth values. The increasing zero mean Gaussian noise energy applied to the state vector parameters has a minimum standard deviation of

0.625% and maximum of 10% (see table 6.1).

For the second evaluation, we used the same incident rays which were perturbed with noise. The initial estimates were then computed using our presented algorithm. The second step was thus performed by starting the iterations from these computed initial estimates.

The results for the first evaluation are presented in the set of figures 6.10 and 6.11. The standard deviation of the zero mean Gaussian noise applied to the pixel coordinates and to the state vector parameters are both expressed in percentage of the ground truth value and increased simultaneously. The used values are presented in table 6.1.

The results for the second evaluation, that is, by estimating the initial state vector using our algorithm, are presented in the set of figures 6.12 and 6.13. The standard deviation of the zero mean Gaussian noise added to the pixel coordinates is the same applied in the previous evaluation and presented in the first row of table 6.1. They are both expressed in percentage of the ground truth value.

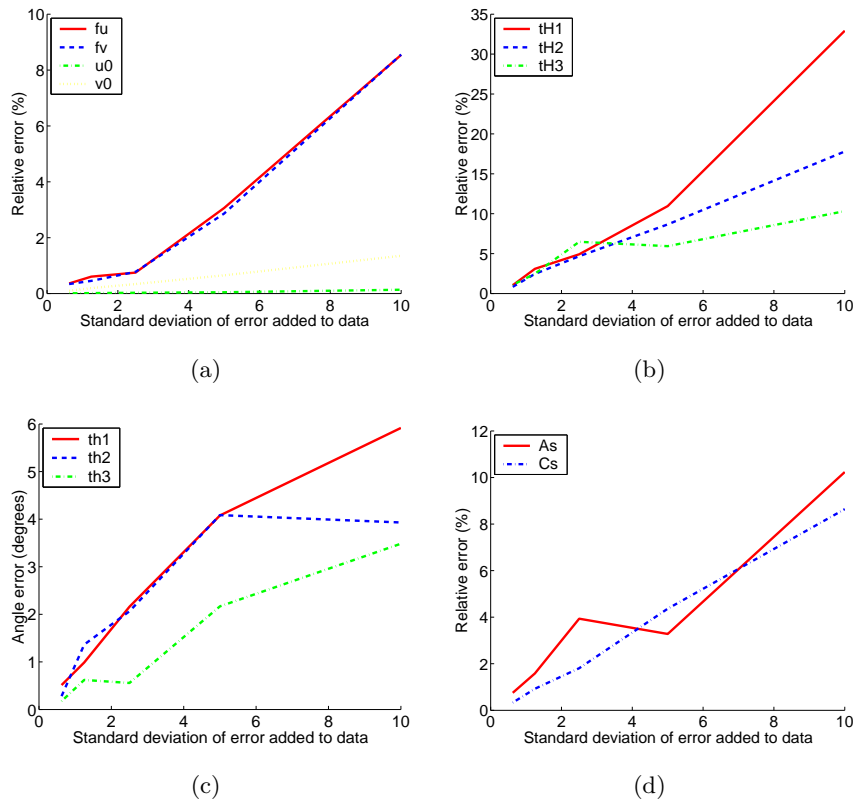


Figure 6.10: Results with **simulated data**. Relative error, or angular error for Euler angles, of the estimated state vector elements, for increasing noise energy added to the initial estimate and to the image pixel coordinates. The method is tested as a whole. The results are plotted in percentage - % for non angular parameters and degrees for angular ones. The parameters whose ground truth value is zero are omitted since no relative error can be computed - ν , θ_2 and B_s . The standard deviation of the noise added is expressed in percentage of the ground truth value.

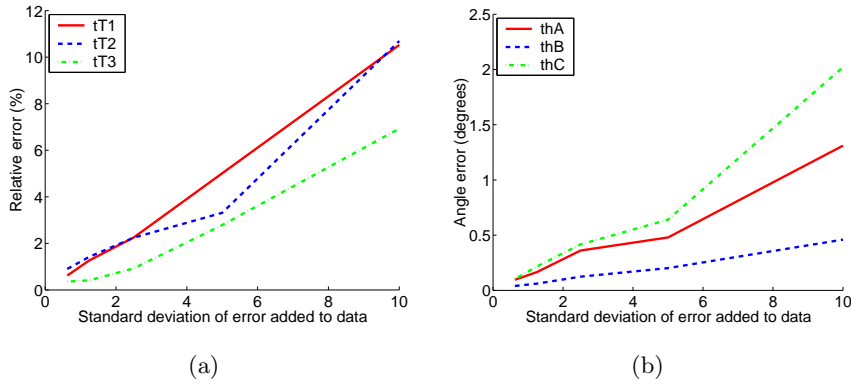


Figure 6.11: **Cont.** - Results with **simulated data**. Relative error, or angular error for Euler angles, of the estimated state vector elements, for increasing noise energy added to the initial estimate and to the image pixel coordinates. The method is tested as a whole. The results are plotted in percentage - % for non angular parameters and degrees for angular ones. The parameters whose ground truth value is zero are omitted since no relative error can be computed - ν , θ_2 and B_s . The standard deviation of the noise added is expressed in percentage of the ground truth value.

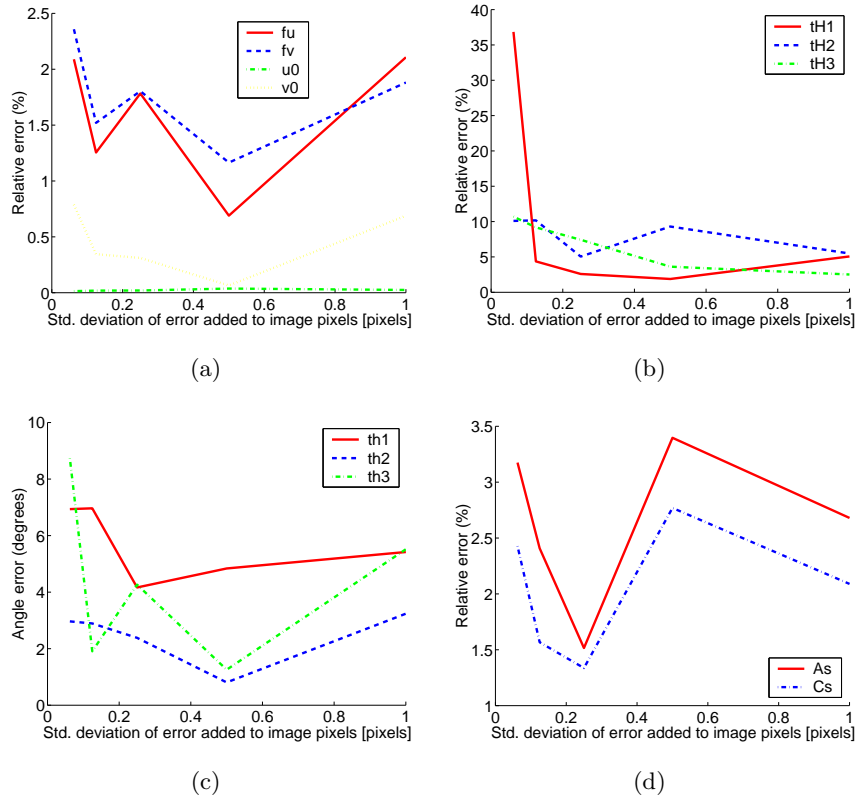


Figure 6.12: Results with **simulated data**. Relative error, or angular error for Euler angles, of the estimated state vector elements, for increasing noise energy added to the image pixel coordinates. The initial estimate is obtained by using our automatic algorithm. The method is tested as a whole. The results are plotted in percentage - % for non angular parameters and degrees for angular ones. The parameters whose ground truth value is zero are omitted since no relative error can be computed - ν , θ_2 and B_s . The standard deviation of the noise added is expressed in percentage of the ground truth value.

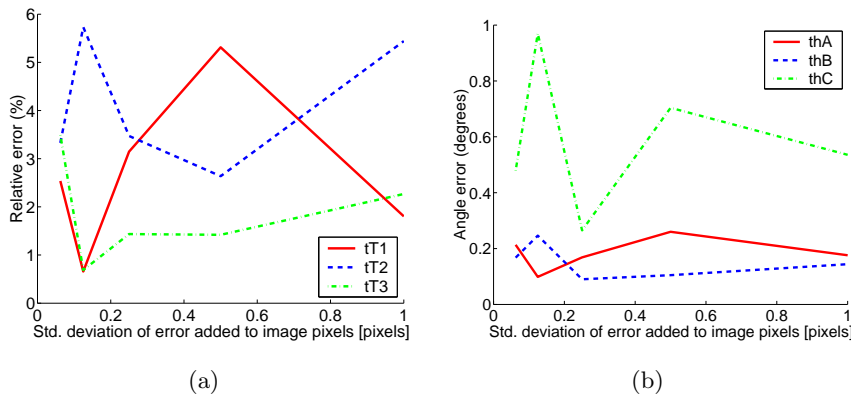


Figure 6.13: **Cont.** - Results with **simulated data**. Relative error, or angular error for Euler angles, of the estimated state vector elements, for increasing noise energy added to the image pixel coordinates. The initial estimate is obtained by using our automatic algorithm. The method is tested as a whole. The results are plotted in percentage - % for non angular parameters and degrees for angular ones. The parameters whose ground truth value is zero are omitted since no relative error can be computed - ν , θ_2 and B_s . The standard deviation of the noise added is expressed in percentage of the ground truth value.

In figures 6.10 to 6.13, for the parameters whose ground truth value is zero, we omitted their estimated values since no relative error can be computed. These parameters are ν , θ_2 and B_s . We observed however, as explained above, that their estimated values were around zero and that their absolute value drifted from zero as the error energy increases, as expected. We also present in figure 6.14 the RMS value of the reprojection error and angle deviation of incident lines, correspondent to the tests using our automatic algorithm for the initial estimate that are presented in figures 6.12 and 6.13. To compute these error measures we did not use the points over which we iterated the bundle adjustment but rather used the remaining point of

the calibration patches.

It can be concluded from the observation of figures 6.10 to 6.14 that the method as a whole is sensitive to noise both in the image pixels and in the initial estimate parameters. Accurate estimation of the vision system parameters (and therefore small reprojection error and deviation from Snell's Law) is generally obtained for small error energies (standard deviation smaller than 1% of the ground truth value). It was also observed from figure 6.10 to 6.13 that the error increased drastically when the first calibration step was introduced, the ray calibration. We conclude that the first step calibration is critical to the final estimation results and a strong effort must be put on it.

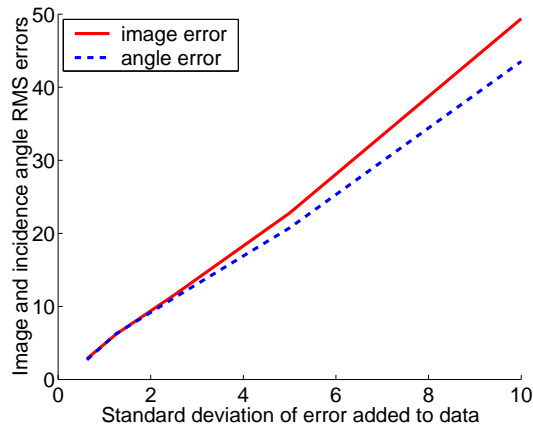


Figure 6.14: Results with **simulated data**. Image error and angle deviation from Snell's Law obtained with the method as a whole, introducing error in the image pixel coordinates and in the incident rays, correspondent to the tests using our automatic algorithm for the initial estimate that are presented in the set of figures 6.12 and 6.13. These error measures were computed in points not used in the bundle adjustment process. The standard deviation of the noise added is expressed in percentage of the ground truth value.

6.5.2 Real Image Experiments

In experiments with real images, we used a commercial high resolution camera CANON EOS 350D (the image is 3456x2304 pixels) with two different mirrors: a spherical and a hyperbolic one. The camera was positioned to be non aligned with the mirror axis in the hyperbolic case and so that its optical axis does not pass through the center of the spherical mirror. These vision system configurations are consequently noncentral.

In order to have ground truth values to compare with, the system was calibrated by applying firstly the Camera Calibration Toolbox [18] to the pinhole camera. The calibration of the mirror and the poses of the camera in relation to the mirror and world reference frames were achieved using our algorithm presented in chapter 5 and in [54]. Some heuristics were also used to find the best solution to the problem. The calibration parameters were also estimated using three images taken with the system in three different positions affected by known motion transformations. Besides the reprojection error, the error measure used in the estimation process was the distance from the 3D world point to the back projected incident ray. In this pre-calibration we achieved a mean distance of $0.6mm$ for the hyperbolic mirror and $2.0mm$ for the spherical one, in a range of approximately $400mm$. Since both the reprojection error and the distance from the 3D points to the back-projected incident rays is small, we consider the pre-calibration results as ground truth to compare with in our experiments.

Figure 6.15 shows two sample images taken by our experimental setup in its hyperbolic and spherical configurations.

Points from the calibration pattern were acquired. The first step ray calibration was computed using four points to estimate the plane homography and then to estimate the incident lines in space. The first estimate algorithm was then run and the second step calibration algorithm iterated until

Table 6.2: Final calibration and intermediate initial estimates relative errors in percentage for the **hyperbolic** configuration. Absolute values are presented for the zero-valued truth parameters.

	True value	Initial estimate	Error (%)	Final calibration	Error (%)
f_u	8086.75	7679.73	5.03	7640.57	5.52
f_v	8058.73	5965.97	25.97	7538.68	6.45
ν	0.00	0.00	-	-0.00	-
u_0	1672.38	2172.30	29.89	1671.67	0.04
v_0	1146.18	1200.27	4.72	1147.95	0.15
t_{H1}	-82.77	-106.02	28.09	-65.55	20.80
t_{H2}	126.94	143.66	13.18	145.18	14.37
t_{H3}	173.20	84.64	51.13	172.38	0.47
θ_1	-4.15	-4.60	10.70	-4.32	3.99
θ_2	0.04	0.07	64.05	0.05	5.83
θ_3	-6.27	-5.72	8.78	-6.73	7.42
A_s	-0.76	-0.95	23.00	-0.92	20.79
B_s	0.00	0.00	-	0.00	-
C_s	-559.96	-590.00	5.64	-537.25	4.06
t_{T1}	-0.57	-0.45	20.71	-0.51	9.97
t_{T2}	4.79	4.44	7.30	4.74	1.02
t_{T3}	292.76	371.34	26.84	270.70	7.54
θ_A	0.19	0.23	20.18	0.18	3.03
θ_B	0.14	0.15	4.28	0.15	2.98
θ_C	-0.41	-0.43	5.99	-0.40	1.37

Table 6.3: Final calibration and intermediate initial estimates relative errors in percentage for the **spherical** configuration. Absolute values are presented for the zero-valued truth parameters.

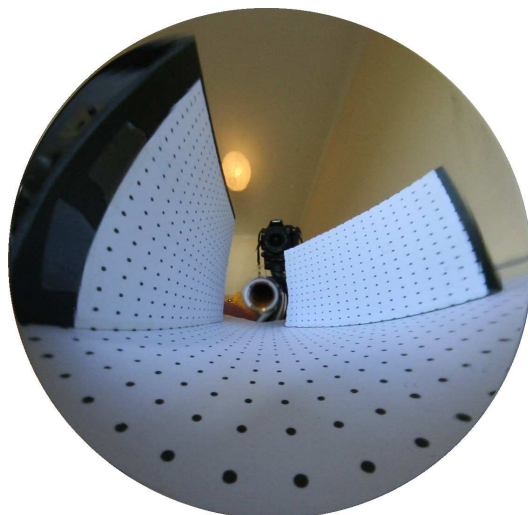
	True value	Initial estimate	Error (%)	Final calibration	Error (%)
f_u	7702.53	5724.58	25.68	7789.50	1.13
f_v	7601.90	5649.66	25.68	7687.46	1.13
ν	0.00	-0.42	-	0.01	-
u_0	1773.37	1685.33	4.96	1771.43	0.11
v_0	1316.00	1217.59	7.48	1319.51	0.27
t_{H1}	-61.06	4.89	108.01	41.05	167.23
t_{H2}	174.73	157.05	10.12	174.31	0.24
t_{H3}	-366.43	-382.83	4.48	-364.35	0.57
θ_1	2.42	2.15	11.45	2.41	0.72
θ_2	0.11	0.07	39.07	0.08	32.60
θ_3	0.14	-0.25	282.82	-0.29	309.68
A_s	1.00	1.30	29.87	0.99	1.36
B_s	0.00	-0.00	-	-0.00	-
C_s	625.00	859.00	37.45	620.97	0.64
t_{T1}	14.59	13.29	8.89	14.47	0.81
t_{T2}	10.48	7.94	24.25	10.57	0.92
t_{T3}	-285.05	-248.69	12.75	-287.32	0.80
θ_A	0.00	0.05	-	-0.00	-
θ_B	0.00	0.00	-	0.00	-
θ_C	0.00	0.01	-	0.62	-

convergence. Tables 6.5.2 and 6.3 show the results and the relative error for each state vector component, for the hyperbolic and spherical configurations respectively.

From the results one can conclude that the first estimate permitted the algorithm to converge to an acceptable solution.



(a) Spherical mirror



(b) Hyperbolic mirror

Figure 6.15: Real images taken with our experimental setup.

6.6 Summary and Conclusions

This chapter describes a two-steps method to estimate the parameters of general catadioptric vision systems.

The first step is the ray calibration in such a way that it provides correspondences between pixels and incident lines in space, in a general and unconstrained world reference frame. The second step is the estimation of the intrinsic and extrinsic parameters of the camera (considered to be a pin-hole) in relation to the mirror surface and the world reference frame. The mirror shape parameters are also calibrated.

In order to simplify the projection and to provide the method with explicit expressions of the projection and its Jacobian, a parameterized projection model relaxing Snell's Law is derived. The reflection point is considered to be the intersection between the incident line in space and the quadric mirror surface. The intersection point is projected into the image according to the camera model. The parameterized projection model relates the coordinates of the point in the image with the incident rays.

A bundle adjustment method is applied to this model and to the data available in order to iterate the values of the state vector made up by the system parameters - pinhole intrinsic parameters, position and orientation of the camera in the world coordinate system and the mirror shape parameters. The computation of the initial estimate and of the Jacobian, both required by the method, are also addressed.

In relation to the gain coefficients W_C , W_I and W_A , we noticed that they always influence highly the convergence since they weight the components of the cost function: reprojection error (W_I), Snell's Law deviation (W_C) and apparent contour deviation (W_A). We noticed in the course of the experiments that the coefficients of the reprojection error and Snell's Law deviation have similar influence and then must be tuned to have comparable

magnitudes. On the other hand, the apparent contour, if used, has a high influence in the cost function and dramatically enhance the accuracy of the estimation. The value of its coefficient must then be tuned in order to guarantee that its term is of the same magnitude order than the other two terms. Notice, however, that this tuning is highly variable and always depends from images to images.

General results of the experiments with simulated data and real images with different mirror configurations showed that the method is accurate and in general converges to the global minimum or at least to a local minimum with very low value of the cost function. The method is, however, very sensitive to noise and experiments suggest that a great effort must be put in the calibration of the first step, that is, the ray calibration is critical to the accuracy of the estimated parameters.

Appendix A: Jacobian of Projection Model

The explicit extended expression for the Jacobian of the cost function, sum of equations (6.15) and (6.16), are here derived and presented.

The first component of the Jacobian is \mathbf{J}_I , given by:

$$\mathbf{J}_I = W_I \begin{bmatrix} 2(u - u^{measure}) \frac{\partial u}{\partial f_u} + 2(v - v^{measure}) \frac{\partial v}{\partial f_u} \\ 2(u - u^{measure}) \frac{\partial u}{\partial f_v} + 2(v - v^{measure}) \frac{\partial v}{\partial f_v} \\ 2(u - u^{measure}) \frac{\partial u}{\partial v} + 2(v - v^{measure}) \frac{\partial v}{\partial v} \\ 2(u - u^{measure}) \frac{\partial u}{\partial u_0} + 2(v - v^{measure}) \frac{\partial v}{\partial u_0} \\ 2(u - u^{measure}) \frac{\partial u}{\partial v_0} + 2(v - v^{measure}) \frac{\partial v}{\partial v_0} \\ 2(u - u^{measure}) \frac{\partial u}{\partial t_{H1}} + 2(v - v^{measure}) \frac{\partial v}{\partial t_{H1}} \\ \dots \\ 2(u - u^{measure}) \frac{\partial u}{\partial \theta_C} + 2(v - v^{measure}) \frac{\partial v}{\partial \theta_C} \end{bmatrix} \quad (6.22)$$

Taking derivatives of the equations for u and v (the image coordinates of the points), the equations for the Jacobian are a function of the derivatives of \mathbf{R}_{CAM} with respect to each of the components of the state vector \mathbf{x} . However, before computing the derivatives in relation to the state vector parameters, we first provide closed-form equations to the projection model (the equations are closed-form up to a root selection).

The image coordinates are given by :

$$\begin{cases} u = \frac{f_u R_{CAM1} + v R_{CAM2} + u_0 R_{CAM3}}{R_{CAM3}} \\ v = \frac{f_v R_{CAM2} + v_0 R_{CAM3}}{R_{CAM3}} \end{cases} \quad (6.23)$$

and the explicit expressions of the reflection point \mathbf{R}_{CAM} are written as:

$$\begin{bmatrix} \mathbf{R}_{CAM} \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{A} + \alpha(\mathbf{B} - \mathbf{A}) \\ 1 \end{bmatrix} \quad (6.24)$$

with:

$$\begin{aligned}
h_{11} &= h_{11}^f C\theta_2 C\theta_3 + h_{12}^f C\theta_2 S\theta_3 - h_{13}^f S\theta_2 \\
h_{12} &= h_{11}^f (S\theta_1 S\theta_2 C\theta_3 - C\theta_1 S\theta_3) + h_{12}^f (S\theta_1 S\theta_2 S\theta_3 + C\theta_1 C\theta_3) + \\
&\quad + h_{13}^f S\theta_1 C\theta_2 \\
h_{13} &= h_{11}^f (C\theta_1 S\theta_2 C\theta_3 + S\theta_1 S\theta_3) + h_{12}^f (C\theta_1 S\theta_2 S\theta_3 - S\theta_1 C\theta_3) + \\
&\quad + h_{13}^f C\theta_1 C\theta_2 \\
h_{21} &= h_{21}^f C\theta_2 C\theta_3 + h_{22}^f C\theta_2 S\theta_3 - h_{23}^f S\theta_2 \\
h_{22} &= h_{21}^f (S\theta_1 S\theta_2 C\theta_3 - C\theta_1 S\theta_3) + h_{22}^f (S\theta_1 S\theta_2 S\theta_3 + C\theta_1 C\theta_3) + \\
&\quad + h_{23}^f S\theta_1 C\theta_2 \tag{6.27} \\
h_{23} &= h_{21}^f (C\theta_1 S\theta_2 C\theta_3 + S\theta_1 S\theta_3) + h_{22}^f (C\theta_1 S\theta_2 S\theta_3 - S\theta_1 C\theta_3) + \\
&\quad + h_{23}^f C\theta_1 C\theta_2 \\
h_{31} &= h_{31}^f C\theta_2 C\theta_3 + h_{32}^f C\theta_2 S\theta_3 - h_{33}^f S\theta_2 \\
h_{32} &= h_{31}^f (S\theta_1 S\theta_2 C\theta_3 - C\theta_1 S\theta_3) + h_{32}^f (S\theta_1 S\theta_2 S\theta_3 + C\theta_1 C\theta_3) + \\
&\quad + h_{33}^f S\theta_1 C\theta_2 \\
h_{33} &= h_{31}^f (C\theta_1 S\theta_2 C\theta_3 + S\theta_1 S\theta_3) + h_{32}^f (C\theta_1 S\theta_2 S\theta_3 - S\theta_1 C\theta_3) + \\
&\quad + h_{33}^f C\theta_1 C\theta_2
\end{aligned}$$

The expression for the rotation matrix of the camera-mirror pose (given by $\mathbf{R}_{\mathbf{T}} = \{t_{ij}\}$) is written similarly, on the angles θ_A , θ_B and θ_C and as function of the frozen rotation elements t_{ij}^f .

As for the equations of the quadric mirror parameters in camera coordinates it yields $\mathbf{Q}_{\text{CAM}} = \mathbf{T}^{-T} \mathbf{Q} \mathbf{T}^{-1}$. Expanding this equations we obtain the following expressions for the elements of the quadric mirror in camera coordinates:

$$\begin{aligned}
q_{CAM11} &= t_{11}^2 + t_{12}^2 + t_{13}^2 A_s \\
q_{CAM12} &= t_{11}t_{21} + t_{12}t_{22} + t_{13}t_{23} A_s \\
q_{CAM13} &= t_{11}t_{31} + t_{12}t_{32} + t_{13}t_{33} A_s \\
q_{CAM14} &= -t_{11}^2 t_{T1} - t_{11}t_{21}t_{T2} - t_{11}t_{31}t_{T3} - t_{12}^2 t_{T1} - t_{12}t_{22}t_{T2} - \\
&\quad -t_{12}t_{32}t_{T3} - t_{13}^2 t_{T1} A_s - t_{13}t_{23}t_{T2} A_s - t_{13}t_{33}t_{T3} A_s + t_{13} B_s \\
q_{CAM22} &= t_{21}^2 + t_{22}^2 + t_{23}^2 A_s \\
q_{CAM23} &= t_{21}t_{31} + t_{22}t_{32} + t_{23}t_{33} A_s \\
q_{CAM24} &= -t_{11}t_{21}t_{T1} - t_{21}^2 t_{T2} - t_{21}t_{31}t_{T3} - t_{12}t_{22}t_{T1} - t_{22}^2 t_{T2} - \\
&\quad -t_{22}t_{32}t_{T3} - t_{13}t_{23}t_{T1} A_s - t_{23}^2 t_{T2} A_s - t_{23}t_{33}t_{T3} A_s + t_{23} B_s \\
q_{CAM33} &= t_{31}^2 + t_{32}^2 + t_{33}^2 A_s \\
q_{CAM34} &= -t_{11}t_{31}t_{T1} - t_{21}t_{31}t_{T2} - t_{31}^2 t_{T3} - t_{12}t_{32}t_{T1} - t_{22}t_{32}t_{T2} - \\
&\quad -t_{32}^2 t_{T3} - t_{13}t_{33}t_{T1} A_s - t_{23}t_{33}t_{T2} A_s - t_{33}^2 t_{T3} A_s + t_{33} B_s \\
q_{CAM44} &= (t_{11}t_{T1} + t_{21}t_{T2} + t_{31}t_{T3})^2 + (t_{12}t_{T1} + t_{22}t_{T2} + t_{32}t_{T3})^2 + \\
&\quad + (t_{13}t_{T1} + t_{23}t_{T2} + t_{33}t_{T3})^2 A_s - \\
&\quad -2(t_{13}t_{T1} + t_{23}t_{T2} + t_{33}t_{T3}) B_s - C_s
\end{aligned} \tag{6.28}$$

where only the elements of the upper triangular matrix are presented since the quadric matrix is symmetric.

In order to close the expressions of the projection model, one has to express α as function of the state vector parameters. Its explicit expression is the expansion and solution of equation (6.5). It yields:

$$\begin{aligned}
\alpha &= \frac{-C_2 \pm \sqrt{C_2^2 - 4C_1C_3}}{2C_1} & (6.29) \\
C_1 &= q_{CAM11}D_1^2 + 2q_{CAM12}D_1D_2 + 2q_{CAM13}D_1D_3 + q_{CAM22}D_2^2 + \\
&\quad + 2q_{CAM23}D_2D_3 + q_{CAM33}D_3^2 \\
C_2 &= 2q_{CAM11}A_1D_1 + 2q_{CAM12}(A_1D_2 + A_2D_1) + \\
&\quad + 2q_{CAM13}(A_1D_3 + A_3D_1) + 2q_{CAM14}D_1 + 2q_{CAM22}A_2D_2 + \\
&\quad + 2q_{CAM23}(A_2D_3 + A_3D_2) + 2q_{CAM24}D_2 + 2q_{CAM33}A_3D_3 + \\
&\quad + 2q_{CAM34}D_3 \\
C_3 &= q_{CAM11}A_1^2 + 2q_{CAM12}A_1A_2 + 2q_{CAM13}A_1A_3 + 2q_{CAM14}A_1 + \\
&\quad + q_{CAM22}A_2^2 + 2q_{CAM23}A_2A_3 + 2q_{CAM24}A_2 + q_{CAM33}A_3^2 + \\
&\quad + 2q_{CAM34}A_3 + q_{CAM44}
\end{aligned}$$

where A_i is the i -th element of \mathbf{A} , B_i the i -th element of \mathbf{B} and $D_i = B_i - A_i$.

Since α is the solution of a second degree polynomial with two distinct real roots and since there is no immediate way to choose the appropriate one (for which we discuss a strategy in section 6.2), the expressions of the image coordinates (u, v) are not actually closed. However, as discussed in section 6.2, if some simple previous conditions about the relative positions of points \mathbf{A}_w and \mathbf{B}_w are met one knows that the appropriate root is the one with the smallest absolute value, hence completing the expressions of (u, v) and hence of the cost function.

Concerning the derivatives of the image pixel coordinates, we have:

$$\begin{aligned}
\frac{\partial u}{\partial f_u} &= \frac{R_{CAM1}}{R_{CAM3}} \\
\frac{\partial u}{\partial f_v} &= 0 \\
\frac{\partial u}{\partial \nu} &= \frac{R_{CAM2}}{R_{CAM3}}
\end{aligned}$$

$$\begin{aligned} \frac{\partial u}{\partial u_0} &= 1 \\ \frac{\partial u}{\partial v_0} &= 0 \\ \frac{\partial u}{\partial x_k} &= \frac{\left(f_u \frac{\partial R_{CAM_1}}{\partial x_k} + \nu \frac{\partial R_{CAM_2}}{\partial x_k}\right) R_{CAM_3} - (f_u R_{CAM_1} + \nu R_{CAM_2}) \frac{\partial R_{CAM_3}}{\partial x_k}}{R_{CAM_3}^2}, \\ &\text{where } x_k \in \{t_{H1}, t_{H2}, t_{H3}, \theta_1, \theta_2, \theta_3, A_s, \dots, \theta_B, \theta_C\} \\ \frac{\partial v}{\partial f_u} &= 0 \\ \frac{\partial v}{\partial f_v} &= \frac{R_{CAM_2}}{R_{CAM_3}} \\ \frac{\partial v}{\partial \nu} &= 0 \\ \frac{\partial v}{\partial u_0} &= 0 \\ \frac{\partial v}{\partial v_0} &= 1 \\ \frac{\partial v}{\partial x_k} &= \frac{f_v \frac{\partial R_{CAM_2}}{\partial x_k} R_{CAM_3} - f_v R_{CAM_2} \frac{\partial R_{CAM_3}}{\partial x_k}}{R_{CAM_3}^2}, \\ &\text{where } x_k \in \{t_{H1}, t_{H2}, t_{H3}, \theta_1, \theta_2, \theta_3, A_s, \dots, \theta_B, \theta_C\} \end{aligned}$$

Since we apply the chain rule to compute the full derivatives, we start by the expression of the rotation elements h_{ij} and t_{ij} derivatives, then by the expressions of the derivatives of the quadric mirror elements $q_{CAM_{ij}}$, followed by the expressions of the scalar α derivatives and finally by the derivatives of the reflection point coordinates R_{CAM_i} .

Since the rotation elements h_{ij} and t_{ij} have similar form, depending on the triplets $(\theta_1, \theta_2, \theta_3)$ and $(\theta_A, \theta_B, \theta_C)$, respectively, we opt to present only the expressions of the derivatives of h_{ij} . To obtain the explicit derivatives of t_{ij} one just has to interchange the triplets of Euler angles in the corresponding expressions. We hence obtain:

$$\begin{aligned}
\frac{\partial h_{11}}{\partial \theta_1} &= 0 \\
\frac{\partial h_{11}}{\partial \theta_2} &= -h_{11}^f S \theta_2 C \theta_3 - h_{12}^f S \theta_2 S \theta_3 - h_{13}^f C \theta_2 \\
\frac{\partial h_{11}}{\partial \theta_3} &= -h_{11}^f C \theta_2 S \theta_3 + h_{12}^f C \theta_2 C \theta_3
\end{aligned} \tag{6.30}$$

$$\begin{aligned}
\frac{\partial h_{12}}{\partial \theta_1} &= h_{11}^f (C \theta_1 S \theta_2 C \theta_3 + S \theta_1 S \theta_3) + h_{12}^f (C \theta_1 S \theta_2 S \theta_3 - S \theta_1 C \theta_3) + \\
&\quad + h_{13}^f C \theta_1 C \theta_2 \\
\frac{\partial h_{12}}{\partial \theta_2} &= h_{11}^f S \theta_1 C \theta_2 C \theta_3 + h_{12}^f S \theta_1 C \theta_2 S \theta_3 - h_{13}^f S \theta_1 S \theta_2 \\
\frac{\partial h_{12}}{\partial \theta_3} &= -h_{11}^f (S \theta_1 S \theta_2 S \theta_3 + C \theta_1 C \theta_3) + h_{12}^f (S \theta_1 S \theta_2 C \theta_3 - C \theta_1 S \theta_3)
\end{aligned} \tag{6.31}$$

$$\begin{aligned}
\frac{\partial h_{13}}{\partial \theta_1} &= h_{11}^f (C \theta_1 S \theta_3 - S \theta_1 S \theta_2 C \theta_3) - h_{12}^f (S \theta_1 S \theta_2 S \theta_3 + C \theta_1 C \theta_3) - \\
&\quad - h_{13}^f S \theta_1 C \theta_2 \\
\frac{\partial h_{13}}{\partial \theta_2} &= h_{11}^f C \theta_1 C \theta_2 C \theta_3 + h_{12}^f C \theta_1 C \theta_2 S \theta_3 - h_{13}^f C \theta_1 S \theta_2 \\
\frac{\partial h_{13}}{\partial \theta_3} &= h_{11}^f (S \theta_1 C \theta_3 - C \theta_1 S \theta_2 S \theta_3) + h_{12}^f (C \theta_1 S \theta_2 C \theta_3 + S \theta_1 S \theta_3)
\end{aligned} \tag{6.32}$$

$$\begin{aligned}
\frac{\partial h_{21}}{\partial \theta_1} &= 0 \\
\frac{\partial h_{21}}{\partial \theta_2} &= -h_{21}^f S \theta_2 C \theta_3 - h_{22}^f S \theta_2 S \theta_3 - h_{23}^f C \theta_2 \\
\frac{\partial h_{21}}{\partial \theta_3} &= -h_{21}^f C \theta_2 S \theta_3 + h_{22}^f C \theta_2 C \theta_3
\end{aligned} \tag{6.33}$$

$$\begin{aligned}
\frac{\partial h_{22}}{\partial \theta_1} &= h_{21}^f (C \theta_1 S \theta_2 C \theta_3 + S \theta_1 S \theta_3) + h_{22}^f (C \theta_1 S \theta_2 S \theta_3 - S \theta_1 C \theta_3) + \\
&\quad + h_{23}^f C \theta_1 C \theta_2 \\
\frac{\partial h_{22}}{\partial \theta_2} &= h_{21}^f S \theta_1 C \theta_2 C \theta_3 + h_{22}^f S \theta_1 C \theta_2 S \theta_3 - h_{23}^f S \theta_1 S \theta_2 \\
\frac{\partial h_{22}}{\partial \theta_3} &= -h_{21}^f (S \theta_1 S \theta_2 S \theta_3 + C \theta_1 C \theta_3) + h_{22}^f (S \theta_1 S \theta_2 C \theta_3 - C \theta_1 S \theta_3)
\end{aligned} \tag{6.34}$$

$$\begin{aligned}
\frac{\partial h_{23}}{\partial \theta_1} &= h_{21}^f (C\theta_1 S\theta_3 - S\theta_1 S\theta_2 C\theta_3) - h_{22}^f (S\theta_1 S\theta_2 S\theta_3 + C\theta_1 C\theta_3) - \\
&\quad - h_{23}^f S\theta_1 C\theta_2 \\
\frac{\partial h_{23}}{\partial \theta_2} &= h_{21}^f C\theta_1 C\theta_2 C\theta_3 + h_{22}^f C\theta_1 C\theta_2 S\theta_3 - h_{23}^f C\theta_1 S\theta_2 \\
\frac{\partial h_{23}}{\partial \theta_3} &= h_{21}^f (S\theta_1 C\theta_3 - C\theta_1 S\theta_2 S\theta_3) + h_{22}^f (C\theta_1 S\theta_2 C\theta_3 + S\theta_1 S\theta_3)
\end{aligned} \tag{6.35}$$

$$\begin{aligned}
\frac{\partial h_{31}}{\partial \theta_1} &= 0 \\
\frac{\partial h_{31}}{\partial \theta_2} &= -h_{31}^f S\theta_2 C\theta_3 - h_{32}^f S\theta_2 S\theta_3 - h_{33}^f C\theta_2 \\
\frac{\partial h_{31}}{\partial \theta_3} &= -h_{31}^f C\theta_2 S\theta_3 + h_{32}^f C\theta_2 C\theta_3
\end{aligned} \tag{6.36}$$

$$\begin{aligned}
\frac{\partial h_{32}}{\partial \theta_1} &= h_{31}^f (C\theta_1 S\theta_2 C\theta_3 + S\theta_1 S\theta_3) + h_{32}^f (C\theta_1 S\theta_2 S\theta_3 - S\theta_1 C\theta_3) + \\
&\quad + h_{33}^f C\theta_1 C\theta_2 \\
\frac{\partial h_{32}}{\partial \theta_2} &= h_{31}^f S\theta_1 C\theta_2 C\theta_3 + h_{32}^f S\theta_1 C\theta_2 S\theta_3 - h_{33}^f S\theta_1 S\theta_2 \\
\frac{\partial h_{32}}{\partial \theta_3} &= -h_{31}^f (S\theta_1 S\theta_2 S\theta_3 + C\theta_1 C\theta_3) + h_{32}^f (S\theta_1 S\theta_2 C\theta_3 - C\theta_1 S\theta_3)
\end{aligned} \tag{6.37}$$

$$\begin{aligned}
\frac{\partial h_{33}}{\partial \theta_1} &= h_{31}^f (C\theta_1 S\theta_3 - S\theta_1 S\theta_2 C\theta_3) - h_{32}^f (S\theta_1 S\theta_2 S\theta_3 + C\theta_1 C\theta_3) - \\
&\quad - h_{33}^f S\theta_1 C\theta_2 \\
\frac{\partial h_{33}}{\partial \theta_2} &= h_{31}^f C\theta_1 C\theta_2 C\theta_3 + h_{32}^f C\theta_1 C\theta_2 S\theta_3 - h_{33}^f C\theta_1 S\theta_2 \\
\frac{\partial h_{33}}{\partial \theta_3} &= h_{31}^f (S\theta_1 C\theta_3 - C\theta_1 S\theta_2 S\theta_3) + h_{32}^f (C\theta_1 S\theta_2 C\theta_3 + S\theta_1 S\theta_3)
\end{aligned} \tag{6.38}$$

Since the update rotation angles are around zero and in the beginning of every iteration their values are zero, we substitute it directly in the equations of the derivatives, providing very simple derivative equations as every $S\theta_i$ is 0 and $C\theta_i$ is 1. The final simplified expressions are given by:

$$\begin{array}{lll}
\frac{\partial h_{11}}{\partial \theta_1} = 0 & \frac{\partial h_{11}}{\partial \theta_2} = -h_{13}^f & \frac{\partial h_{11}}{\partial \theta_3} = h_{12}^f \\
\frac{\partial h_{12}}{\partial \theta_1} = h_{13}^f & \frac{\partial h_{12}}{\partial \theta_2} = 0 & \frac{\partial h_{12}}{\partial \theta_3} = -h_{11}^f \\
\frac{\partial h_{13}}{\partial \theta_1} = -h_{12}^f & \frac{\partial h_{13}}{\partial \theta_2} = h_{11}^f & \frac{\partial h_{13}}{\partial \theta_3} = 0 \\
\frac{\partial h_{21}}{\partial \theta_1} = 0 & \frac{\partial h_{21}}{\partial \theta_2} = -h_{23}^f & \frac{\partial h_{21}}{\partial \theta_3} = h_{22}^f \\
\frac{\partial h_{22}}{\partial \theta_1} = h_{23}^f & \frac{\partial h_{22}}{\partial \theta_2} = 0 & \frac{\partial h_{22}}{\partial \theta_3} = -h_{21}^f \\
\frac{\partial h_{23}}{\partial \theta_1} = -h_{22}^f & \frac{\partial h_{23}}{\partial \theta_2} = h_{21}^f & \frac{\partial h_{23}}{\partial \theta_3} = 0 \\
\frac{\partial h_{31}}{\partial \theta_1} = 0 & \frac{\partial h_{31}}{\partial \theta_2} = -h_{33}^f & \frac{\partial h_{31}}{\partial \theta_3} = h_{32}^f \\
\frac{\partial h_{32}}{\partial \theta_1} = h_{33}^f & \frac{\partial h_{32}}{\partial \theta_2} = 0 & \frac{\partial h_{32}}{\partial \theta_3} = -h_{31}^f \\
\frac{\partial h_{33}}{\partial \theta_1} = -h_{32}^f & \frac{\partial h_{33}}{\partial \theta_2} = h_{31}^f & \frac{\partial h_{33}}{\partial \theta_3} = 0
\end{array}$$

and $\partial h_{ij}/\partial x_k = 0$ for all x_k except θ_1 , θ_2 and θ_3 .

Similarly, the derivatives of the camera-mirror pose rotation elements are given by:

$$\begin{array}{lll}
\frac{\partial t_{11}}{\partial \theta_A} = 0 & \frac{\partial t_{11}}{\partial \theta_B} = -t_{13}^f & \frac{\partial t_{11}}{\partial \theta_C} = t_{12}^f \\
\frac{\partial t_{12}}{\partial \theta_A} = t_{13}^f & \frac{\partial t_{12}}{\partial \theta_B} = 0 & \frac{\partial t_{12}}{\partial \theta_C} = -t_{11}^f \\
\frac{\partial t_{13}}{\partial \theta_A} = -t_{12}^f & \frac{\partial t_{13}}{\partial \theta_B} = t_{11}^f & \frac{\partial t_{13}}{\partial \theta_C} = 0 \\
\frac{\partial t_{21}}{\partial \theta_A} = 0 & \frac{\partial t_{21}}{\partial \theta_B} = -t_{23}^f & \frac{\partial t_{21}}{\partial \theta_C} = t_{22}^f \\
\frac{\partial t_{22}}{\partial \theta_A} = t_{23}^f & \frac{\partial t_{22}}{\partial \theta_B} = 0 & \frac{\partial t_{22}}{\partial \theta_C} = -t_{21}^f \\
\frac{\partial t_{23}}{\partial \theta_A} = -t_{22}^f & \frac{\partial t_{23}}{\partial \theta_B} = t_{21}^f & \frac{\partial t_{23}}{\partial \theta_C} = 0 \\
\frac{\partial t_{31}}{\partial \theta_A} = 0 & \frac{\partial t_{31}}{\partial \theta_B} = -t_{33}^f & \frac{\partial t_{31}}{\partial \theta_C} = t_{32}^f \\
\frac{\partial t_{32}}{\partial \theta_A} = t_{33}^f & \frac{\partial t_{32}}{\partial \theta_B} = 0 & \frac{\partial t_{32}}{\partial \theta_C} = -t_{31}^f \\
\frac{\partial t_{33}}{\partial \theta_A} = -t_{32}^f & \frac{\partial t_{33}}{\partial \theta_B} = t_{31}^f & \frac{\partial t_{33}}{\partial \theta_C} = 0
\end{array}$$

and $\partial t_{ij}/\partial x_k = 0$ for all x_k except θ_A , θ_B and θ_C .

If we now derive the quadric mirror elements expressed in equation (6.28), it yields,

$$\begin{aligned}
\frac{\partial q_{CAM_{ij}}}{\partial x_k} &= 0, \text{ for } x_k \in \{f_u, f_v, \nu, u_0, v_0, t_{H1}, t_{H2}, t_{H3}, \theta_1, \theta_2, \theta_3\} \\
&\text{for all } (i, j) \\
\frac{\partial q_{CAM_{11}}}{\partial A_s} &= t_{13}^2 \\
\frac{\partial q_{CAM_{11}}}{\partial \theta_k} &= 2t_{11} \frac{\partial t_{11}}{\partial \theta_k} + 2t_{12} \frac{\partial t_{12}}{\partial \theta_k} + 2A_s t_{13} \frac{\partial t_{13}}{\partial \theta_k} \\
\frac{\partial q_{CAM_{12}}}{\partial A_s} &= t_{13} t_{23} \\
\frac{\partial q_{CAM_{12}}}{\partial \theta_k} &= t_{11} \frac{\partial t_{21}}{\partial \theta_k} + t_{21} \frac{\partial t_{11}}{\partial \theta_k} + t_{12} \frac{\partial t_{22}}{\partial \theta_k} + t_{22} \frac{\partial t_{12}}{\partial \theta_k} + A_s t_{13} \frac{\partial t_{23}}{\partial \theta_k} + \\
&\quad + A_s t_{23} \frac{\partial t_{13}}{\partial \theta_k}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial q_{CAM13}}{\partial A_s} &= t_{13}t_{33} \\
\frac{\partial q_{CAM13}}{\partial \theta_k} &= t_{11}\frac{\partial t_{31}}{\partial \theta_k} + t_{31}\frac{\partial t_{11}}{\partial \theta_k} + t_{12}\frac{\partial t_{32}}{\partial \theta_k} + t_{32}\frac{\partial t_{12}}{\partial \theta_k} + A_s t_{13}\frac{\partial t_{33}}{\partial \theta_k} + \\
&\quad + A_s t_{33}\frac{\partial t_{13}}{\partial \theta_k} \\
\frac{\partial q_{CAM14}}{\partial A_s} &= -t_{13}^2 t_{T1} - t_{13}t_{23}t_{T2} - t_{13}t_{33}t_{T3} \\
\frac{\partial q_{CAM14}}{\partial B_s} &= t_{13} \\
\frac{\partial q_{CAM14}}{\partial t_{T1}} &= -t_{11}^2 - t_{12}^2 - t_{13}^2 A_s \\
\frac{\partial q_{CAM14}}{\partial t_{T2}} &= -t_{11}t_{21} - t_{12}t_{22} - t_{13}t_{23}A_s \\
\frac{\partial q_{CAM14}}{\partial t_{T3}} &= -t_{11}t_{31} - t_{12}t_{32} - t_{13}t_{33}A_s \\
\frac{\partial q_{CAM14}}{\partial \theta_k} &= -2t_{11}t_{T1}\frac{\partial t_{11}}{\partial \theta_k} - t_{11}t_{T2}\frac{\partial t_{21}}{\partial \theta_k} - t_{21}t_{T2}\frac{\partial t_{11}}{\partial \theta_k} - t_{11}t_{T3}\frac{\partial t_{31}}{\partial \theta_k} - \\
&\quad - t_{31}t_{T3}\frac{\partial t_{11}}{\partial \theta_k} - 2t_{12}t_{T1}\frac{\partial t_{12}}{\partial \theta_k} - t_{12}t_{T2}\frac{\partial t_{22}}{\partial \theta_k} - t_{22}t_{T2}\frac{\partial t_{12}}{\partial \theta_k} - \\
&\quad - t_{12}t_{T3}\frac{\partial t_{32}}{\partial \theta_k} - t_{32}t_{T3}\frac{\partial t_{12}}{\partial \theta_k} - 2t_{13}t_{T1}\frac{\partial t_{13}}{\partial \theta_k}A_s - t_{13}t_{T2}\frac{\partial t_{23}}{\partial \theta_k}A_s - \\
&\quad - t_{23}t_{T2}\frac{\partial t_{13}}{\partial \theta_k}A_s - t_{13}t_{T3}\frac{\partial t_{33}}{\partial \theta_k}A_s - t_{33}t_{T3}\frac{\partial t_{13}}{\partial \theta_k}A_s + \frac{\partial t_{13}}{\partial \theta_k}B_s \\
\frac{\partial q_{CAM22}}{\partial A_s} &= t_{23}^2 \\
\frac{\partial q_{CAM22}}{\partial \theta_k} &= 2t_{21}\frac{\partial t_{21}}{\partial \theta_k} + 2t_{22}\frac{\partial t_{22}}{\partial \theta_k} + 2A_s t_{23}\frac{\partial t_{23}}{\partial \theta_k} \\
\frac{\partial q_{CAM23}}{\partial A_s} &= t_{23}t_{33} \\
\frac{\partial q_{CAM23}}{\partial \theta_k} &= t_{21}\frac{\partial t_{31}}{\partial \theta_k} + t_{31}\frac{\partial t_{21}}{\partial \theta_k} + t_{22}\frac{\partial t_{32}}{\partial \theta_k} + t_{32}\frac{\partial t_{22}}{\partial \theta_k} + A_s t_{23}\frac{\partial t_{33}}{\partial \theta_k} + A_s t_{33}\frac{\partial t_{23}}{\partial \theta_k} \\
\frac{\partial q_{CAM24}}{\partial A_s} &= -t_{13}t_{23}t_{T1} - t_{23}^2 t_{T2} - t_{23}t_{33}t_{T3} \\
\frac{\partial q_{CAM24}}{\partial B_s} &= t_{23} \\
\frac{\partial q_{CAM24}}{\partial t_{T1}} &= -t_{11}t_{21} - t_{12}t_{22} - t_{13}t_{23}A_s \\
\frac{\partial q_{CAM24}}{\partial t_{T2}} &= -t_{21}^2 - t_{22}^2 - t_{23}^2 A_s \\
\frac{\partial q_{CAM24}}{\partial t_{T3}} &= -t_{21}t_{31} - t_{22}t_{32} - t_{23}t_{33}A_s
\end{aligned}$$

$$\begin{aligned}
\frac{\partial q_{CAM24}}{\partial \theta_k} &= -t_{11}t_{T1} \frac{\partial t_{21}}{\partial \theta_k} - t_{21}t_{T1} \frac{\partial t_{11}}{\partial \theta_k} - 2t_{21}t_{T2} \frac{\partial t_{21}}{\partial \theta_k} - t_{21}t_{T3} \frac{\partial t_{31}}{\partial \theta_k} - \\
&\quad -t_{31}t_{T3} \frac{\partial t_{21}}{\partial \theta_k} - t_{12}t_{T1} \frac{\partial t_{22}}{\partial \theta_k} - t_{22}t_{T1} \frac{\partial t_{12}}{\partial \theta_k} - 2t_{22}t_{T2} \frac{\partial t_{22}}{\partial \theta_k} - \\
&\quad -t_{22}t_{T3} \frac{\partial t_{32}}{\partial \theta_k} - t_{32}t_{T3} \frac{\partial t_{22}}{\partial \theta_k} - t_{13}t_{T1} \frac{\partial t_{23}}{\partial \theta_k} A_s - t_{23}t_{T1} \frac{\partial t_{13}}{\partial \theta_k} A_s - \\
&\quad -2t_{23}t_{T2} \frac{\partial t_{23}}{\partial \theta_k} A_s - t_{23}t_{T3} \frac{\partial t_{33}}{\partial \theta_k} A_s - t_{33}t_{T3} \frac{\partial t_{23}}{\partial \theta_k} A_s + \frac{\partial t_{23}}{\partial \theta_k} B_s \\
\frac{\partial q_{CAM33}}{\partial A_s} &= t_{33}^2 \\
\frac{\partial q_{CAM33}}{\partial \theta_k} &= 2t_{31} \frac{\partial t_{31}}{\partial \theta_k} + 2t_{32} \frac{\partial t_{32}}{\partial \theta_k} + 2A_s t_{33} \frac{\partial t_{33}}{\partial \theta_k} \\
\frac{\partial q_{CAM34}}{\partial A_s} &= -t_{13}t_{33}t_{T1} - t_{23}t_{33}t_{T2} - t_{33}^2 t_{T3} \\
\frac{\partial q_{CAM34}}{\partial B_s} &= t_{33} \\
\frac{\partial q_{CAM34}}{\partial t_{T1}} &= -t_{11}t_{31} - t_{12}t_{32} - t_{13}t_{33}A_s \\
\frac{\partial q_{CAM34}}{\partial t_{T2}} &= -t_{21}t_{31} - t_{22}t_{32} - t_{23}t_{33}A_s \\
\frac{\partial q_{CAM34}}{\partial t_{T3}} &= -t_{31}^2 - t_{32}^2 - t_{33}^2 A_s \\
\frac{\partial q_{CAM34}}{\partial \theta_k} &= -t_{11}t_{T1} \frac{\partial t_{31}}{\partial \theta_k} - t_{31}t_{T1} \frac{\partial t_{11}}{\partial \theta_k} - t_{21}t_{T2} \frac{\partial t_{31}}{\partial \theta_k} - t_{31}t_{T2} \frac{\partial t_{21}}{\partial \theta_k} - \\
&\quad -2t_{31}t_{T3} \frac{\partial t_{31}}{\partial \theta_k} - t_{12}t_{T1} \frac{\partial t_{32}}{\partial \theta_k} - t_{32}t_{T1} \frac{\partial t_{12}}{\partial \theta_k} - t_{22}t_{T2} \frac{\partial t_{32}}{\partial \theta_k} - \\
&\quad -t_{32}t_{T2} \frac{\partial t_{22}}{\partial \theta_k} - 2t_{32}t_{T3} \frac{\partial t_{32}}{\partial \theta_k} - t_{13}t_{T1} \frac{\partial t_{33}}{\partial \theta_k} A_s - t_{33}t_{T1} \frac{\partial t_{13}}{\partial \theta_k} A_s - \\
&\quad -t_{23}t_{T2} \frac{\partial t_{33}}{\partial \theta_k} A_s - t_{33}t_{T2} \frac{\partial t_{23}}{\partial \theta_k} A_s - 2t_{33}t_{T3} \frac{\partial t_{33}}{\partial \theta_k} A_s + \frac{\partial t_{33}}{\partial \theta_k} B_s \\
\frac{\partial q_{CAM44}}{\partial A_s} &= (t_{13}t_{T1} + t_{23}t_{T2} - +t_{33}t_{T3})^2 \\
\frac{\partial q_{CAM44}}{\partial B_s} &= -2(t_{13}t_{T1} + t_{23}t_{T2} - +t_{33}t_{T3}) \\
\frac{\partial q_{CAM44}}{\partial C_s} &= -1 \\
\frac{\partial q_{CAM44}}{\partial t_{T1}} &= 2(t_{11}t_{T1} + t_{21}t_{T2} + t_{31}t_{T3})t_{11} + 2(t_{12}t_{T1} + t_{22}t_{T2} + t_{32}t_{T3})t_{12} + \\
&\quad + 2(t_{13}t_{T1} + t_{23}t_{T2} + t_{33}t_{T3})t_{13}A_s - 2t_{13}B_s
\end{aligned}$$

$$\begin{aligned} \frac{\partial q_{CAM44}}{\partial t_{T_2}} = & 2(t_{11}t_{T_1} + t_{21}t_{T_2} + t_{31}t_{T_3})t_{21} + 2(t_{12}t_{T_1} + t_{22}t_{T_2} + t_{32}t_{T_3})t_{22} + \\ & + 2(t_{13}t_{T_1} + t_{23}t_{T_2} + t_{33}t_{T_3})t_{23}A_s - 2t_{23}B_s \end{aligned}$$

$$\begin{aligned} \frac{\partial q_{CAM44}}{\partial t_{T_1}} = & 2(t_{11}t_{T_1} + t_{21}t_{T_2} + t_{31}t_{T_3})t_{31} + 2(t_{12}t_{T_1} + t_{22}t_{T_2} + t_{32}t_{T_3})t_{32} + \\ & + 2(t_{13}t_{T_1} + t_{23}t_{T_2} + t_{33}t_{T_3})t_{33}A_s - 2t_{33}B_s \end{aligned}$$

$$\begin{aligned} \frac{\partial q_{CAM44}}{\partial \theta_k} = & 2(t_{11}t_{T_1} + t_{21}t_{T_2} + t_{31}t_{T_3}) \cdot \left(\frac{\partial t_{11}}{\partial \theta_k} t_{T_1} + \frac{\partial t_{21}}{\partial \theta_k} t_{T_2} + \frac{\partial t_{31}}{\partial \theta_k} t_{T_3} \right) + \\ & + 2(t_{12}t_{T_1} + t_{22}t_{T_2} + t_{32}t_{T_3}) \cdot \left(\frac{\partial t_{12}}{\partial \theta_k} t_{T_1} + \frac{\partial t_{22}}{\partial \theta_k} t_{T_2} + \frac{\partial t_{32}}{\partial \theta_k} t_{T_3} \right) + \\ & + 2(t_{13}t_{T_1} + t_{23}t_{T_2} + t_{33}t_{T_3}) \cdot \left(\frac{\partial t_{13}}{\partial \theta_k} t_{T_1} + \frac{\partial t_{23}}{\partial \theta_k} t_{T_2} + \frac{\partial t_{33}}{\partial \theta_k} t_{T_3} \right) A_s - \\ & - 2 \left(\frac{\partial t_{13}}{\partial \theta_k} t_{T_1} + \frac{\partial t_{23}}{\partial \theta_k} t_{T_2} + \frac{\partial t_{33}}{\partial \theta_k} t_{T_3} \right) B_s \end{aligned}$$

where $\theta_k \in \{\theta_A, \theta_B, \theta_C\}$.

Concerning the scale factor α its derivatives are given by:

$$\frac{\partial \alpha}{\partial x_k} = 0, \text{ where } x_k \in \{f_u, f_v, \nu, u_0, v_0, t_{H1}, t_{H2}, t_{H3}, \theta_1, \theta_2, \theta_3\}$$

$$\begin{aligned} \frac{\partial \alpha}{\partial x_k} = & \frac{\partial \alpha}{\partial C_1} \sum \left(\frac{\partial C_1}{\partial q_{CAMij}} \frac{\partial q_{CAMij}}{\partial x_k} \right) + \frac{\partial \alpha}{\partial C_2} \sum \left(\frac{\partial C_2}{\partial q_{CAMij}} \frac{\partial q_{CAMij}}{\partial x_k} \right) + \\ & + \frac{\partial \alpha}{\partial C_3} \sum \left(\frac{\partial C_3}{\partial q_{CAMij}} \frac{\partial q_{CAMij}}{\partial x_k} \right), \end{aligned}$$

where $x_k \in \{A_s, B_s, C_s, t_{T_1}, t_{T_2}, t_{T_3}, \theta_A, \theta_B, \theta_C\}$ and

$(i, j) \in \{(1, 1), (1, 2), (1, 3), (1, 4), (2, 2), (2, 3), (2, 4), (3, 3), (3, 4), (4, 4)\}$

with

$$\begin{array}{lll}
\frac{\partial C_1}{\partial q_{CAM_{11}}} = D_1^2 & \frac{\partial C_2}{\partial q_{CAM_{11}}} = 2A_1D_1 & \frac{\partial C_3}{\partial q_{CAM_{11}}} = A_1^2 \\
\frac{\partial C_1}{\partial q_{CAM_{12}}} = 2D_1D_2 & \frac{\partial C_2}{\partial q_{CAM_{12}}} = 2A_1D_2 + 2A_2D_1 & \frac{\partial C_3}{\partial q_{CAM_{12}}} = 2A_1A_2 \\
\frac{\partial C_1}{\partial q_{CAM_{13}}} = 2D_1D_3 & \frac{\partial C_2}{\partial q_{CAM_{13}}} = 2A_1D_3 + 2A_3D_1 & \frac{\partial C_3}{\partial q_{CAM_{13}}} = 2A_1A_3 \\
\frac{\partial C_1}{\partial q_{CAM_{14}}} = 0 & \frac{\partial C_2}{\partial q_{CAM_{14}}} = 2D_1 & \frac{\partial C_3}{\partial q_{CAM_{14}}} = 2A_1 \\
\frac{\partial C_1}{\partial q_{CAM_{22}}} = D_2^2 & \frac{\partial C_2}{\partial q_{CAM_{22}}} = 2A_2D_2 & \frac{\partial C_3}{\partial q_{CAM_{22}}} = A_2^2 \\
\frac{\partial C_1}{\partial q_{CAM_{23}}} = 2D_2D_3 & \frac{\partial C_2}{\partial q_{CAM_{23}}} = 2A_2D_3 + 2A_3D_2 & \frac{\partial C_3}{\partial q_{CAM_{23}}} = 2A_2A_3 \\
\frac{\partial C_1}{\partial q_{CAM_{24}}} = 0 & \frac{\partial C_2}{\partial q_{CAM_{24}}} = 2D_2 & \frac{\partial C_3}{\partial q_{CAM_{24}}} = 2A_2 \\
\frac{\partial C_1}{\partial q_{CAM_{33}}} = D_3^2 & \frac{\partial C_2}{\partial q_{CAM_{33}}} = 2A_3D_3 & \frac{\partial C_3}{\partial q_{CAM_{33}}} = A_3^2 \\
\frac{\partial C_1}{\partial q_{CAM_{34}}} = 0 & \frac{\partial C_2}{\partial q_{CAM_{34}}} = 2D_3 & \frac{\partial C_3}{\partial q_{CAM_{34}}} = 2A_3
\end{array}$$

and

$$\begin{aligned}
\frac{\partial \alpha}{\partial C_1} &= \pm \frac{-C_3}{C_1 \sqrt{C_2^2 - 4C_1C_3}} - \frac{-C_2 \pm \sqrt{C_2^2 - 4C_1C_3}}{2C_1^2} \\
\frac{\partial \alpha}{\partial C_2} &= -\frac{1}{2C_1} \pm \frac{C_2}{2C_1 \sqrt{C_2^2 - 4C_1C_3}} \\
\frac{\partial \alpha}{\partial C_3} &= \pm \frac{-1}{\sqrt{C_2^2 - 4C_1C_3}}
\end{aligned}$$

For the derivatives of the reflection point coordinates \mathbf{R}_{CAM} , we obtain the following expressions:

$$\begin{aligned}
\frac{\partial R_{CAM_i}}{\partial x_k} &= 0, \text{ with } x_k \in \{f_u, f_v, \nu, u_0, v_0\} \text{ and } i \in \{1, 2, 3\} \\
\frac{\partial R_{CAM_1}}{\partial t_{H_1}} &= 1 \\
\frac{\partial R_{CAM_1}}{\partial t_{H_2}} &= 0
\end{aligned}$$

$$\begin{aligned}
\frac{\partial R_{CAM1}}{\partial t_{H3}} &= 0 \\
\frac{\partial R_{CAM1}}{\partial \theta_k} &= \frac{\partial h_{11}}{\partial \theta_k} (A_{w1} + \alpha(B_{w1} - A_{w1})) + \frac{\partial h_{12}}{\partial \theta_k} (A_{w2} + \alpha(B_{w2} - A_{w2})) + \\
&\quad + \frac{\partial h_{13}}{\partial \theta_k} (A_{w3} + \alpha(B_{w3} - A_{w3})), \text{ where } \theta_k \in \{\theta_1, \theta_2, \theta_3\} \\
\frac{\partial R_{CAM1}}{\partial x_k} &= (h_{11} (B_{w1} - A_{w1}) + h_{12} (B_{w2} - A_{w2}) + h_{13} (B_{w3} - A_{w3})) \frac{\partial \alpha}{\partial x_k}, \\
&\quad \text{with } x_k \in \{A_s, B_s, C_s, t_{T1}, t_{T2}, t_{T3}, \theta_A, \theta_B, \theta_C\} \\
\frac{\partial R_{CAM2}}{\partial t_{H1}} &= 0 \\
\frac{\partial R_{CAM2}}{\partial t_{H2}} &= 1 \\
\frac{\partial R_{CAM2}}{\partial t_{H3}} &= 0 \\
\frac{\partial R_{CAM2}}{\partial \theta_k} &= \frac{\partial h_{21}}{\partial \theta_k} (A_{w1} + \alpha(B_{w1} - A_{w1})) + \frac{\partial h_{22}}{\partial \theta_k} (A_{w2} + \alpha(B_{w2} - A_{w2})) + \\
&\quad + \frac{\partial h_{23}}{\partial \theta_k} (A_{w3} + \alpha(B_{w3} - A_{w3})), \text{ where } \theta_k \in \{\theta_1, \theta_2, \theta_3\} \\
\frac{\partial R_{CAM2}}{\partial x_k} &= (h_{21} (B_{w1} - A_{w1}) + h_{22} (B_{w2} - A_{w2}) + h_{23} (B_{w3} - A_{w3})) \frac{\partial \alpha}{\partial x_k}, \\
&\quad \text{with } x_k \in \{A_s, B_s, C_s, t_{T1}, t_{T2}, t_{T3}, \theta_A, \theta_B, \theta_C\} \\
\frac{\partial R_{CAM3}}{\partial t_{H1}} &= 0 \\
\frac{\partial R_{CAM3}}{\partial t_{H2}} &= 0 \\
\frac{\partial R_{CAM3}}{\partial t_{H3}} &= 1 \\
\frac{\partial R_{CAM3}}{\partial \theta_k} &= \frac{\partial h_{31}}{\partial \theta_k} (A_{w1} + \alpha(B_{w1} - A_{w1})) + \frac{\partial h_{32}}{\partial \theta_k} (A_{w2} + \alpha(B_{w2} - A_{w2})) + \\
&\quad + \frac{\partial h_{33}}{\partial \theta_k} (A_{w3} + \alpha(B_{w3} - A_{w3})), \text{ where } \theta_k \in \{\theta_1, \theta_2, \theta_3\} \\
\frac{\partial R_{CAM3}}{\partial x_k} &= (h_{31} (B_{w1} - A_{w1}) + h_{32} (B_{w2} - A_{w2}) + h_{33} (B_{w3} - A_{w3})) \frac{\partial \alpha}{\partial x_k}, \\
&\quad \text{with } x_k \in \{A_s, B_s, C_s, t_{T1}, t_{T2}, t_{T3}, \theta_A, \theta_B, \theta_C\}
\end{aligned}$$

Finally, the Jacobian \mathbf{J}_I can then be computed by back-replacing the previous equations such that the explicit expressions can be computed by

equation (6.22). Since these equations are huge, we omit them.

For the Jacobian component \mathbf{J}_A , we first present the equations for the reflected vector \mathbf{V}_r and \mathbf{V}'_r which are given by:

$$\begin{aligned}\mathbf{V}'_r &= -\frac{\mathbf{R}_{CAM}}{\|\mathbf{R}_{CAM}\|} \\ \mathbf{V}_r &= \mathbf{V}_i - 2(\mathbf{V}_i^T \mathbf{N})\mathbf{N} = \mathbf{V}_i - 2W\mathbf{N}\end{aligned}\quad (6.39)$$

where

$$\mathbf{V}_i = \frac{1}{(A_1 - B_1)^2 + (A_2 - B_2)^2 + (A_3 - B_3)^2} \begin{bmatrix} A_1 - B_1 \\ A_2 - B_2 \\ A_3 - B_3 \end{bmatrix} \quad (6.40)$$

$$\mathbf{N} = \frac{1}{\Delta_N} \begin{bmatrix} N_1 \\ N_2 \\ N_3 \end{bmatrix} \quad (6.41)$$

$$N_1 = q_{CAM11}R_{CAM1} + q_{CAM12}R_{CAM2} + q_{CAM13}R_{CAM3} + q_{CAM14}$$

$$N_2 = q_{CAM21}R_{CAM1} + q_{CAM22}R_{CAM2} + q_{CAM23}R_{CAM3} + q_{CAM24}$$

$$N_3 = q_{CAM31}R_{CAM1} + q_{CAM32}R_{CAM2} + q_{CAM33}R_{CAM3} + q_{CAM34}$$

$$\Delta_N = \sqrt{N_1^2 + N_2^2 + N_3^2}$$

$$W = \mathbf{V}_i^T \mathbf{N} = \frac{(A_1 - B_1)N_1 + (A_2 - B_2)N_2 + (A_3 - B_3)N_3}{\Delta_N^2 \sqrt{(A_1 - B_1)^2 + (A_2 - B_2)^2 + (A_3 - B_3)^2}} \quad (6.42)$$

where, since the quadric matrix is symmetric, $q_{CAMij} = q_{CAMji}$.

We now present the derivatives of \mathbf{V}'_r with respect to the state vector components.

$$\begin{aligned} \frac{\partial V'_{rj}}{\partial x_k} = & \frac{\frac{\partial R_{CAMj}}{\partial x_k}}{\sqrt{R_{CAM1}^2 + R_{CAM2}^2 + R_{CAM3}^2}} - \\ & - R_{CAMj} \frac{R_{CAM1} \frac{\partial R_{CAM1}}{\partial x_k} + R_{CAM2} \frac{\partial R_{CAM2}}{\partial x_k} + R_{CAM3} \frac{\partial R_{CAM3}}{\partial x_k}}{\sqrt{(R_{CAM1}^2 + R_{CAM2}^2 + R_{CAM3}^2)^3}} \end{aligned} \quad (6.43)$$

for all x_k in the state vector and $j \in \{1, 2, 3\}$.

Concerning the derivatives of \mathbf{V}_r we start by presenting the derivatives of its components:

$$\frac{\partial V_{ij}}{\partial x_k} = 0, \text{ for all } x_k \text{ in the state vector and } j \in \{1, 2, 3\}$$

and

$$\begin{aligned} \frac{\partial N_j}{\partial x_k} = & 0, \text{ where } x_k \in \{f_u, f_v, \nu, u_0, v_0\} \text{ and } j \in \{1, 2, 3\} \\ \frac{\partial N_j}{\partial x_k} = & q_{CAM1j} \frac{\partial R_{CAM1}}{\partial x_k} + q_{CAM2j} \frac{\partial R_{CAM2}}{\partial x_k} + q_{CAM3j} \frac{\partial R_{CAM3}}{\partial x_k}, \\ & \text{where } x_k \in \{t_{H1}, t_{H2}, t_{H3}, \theta_1, \theta_2, \theta_3\} \text{ and } j \in \{1, 2, 3\} \\ \frac{\partial N_j}{\partial x_k} = & q_{CAMj1} \frac{\partial R_{CAM1}}{\partial x_k} + R_{CAM1} \frac{\partial q_{CAMj1}}{\partial x_k} + q_{CAMj2} \frac{\partial R_{CAM2}}{\partial x_k} + \\ & + R_{CAM2} \frac{\partial q_{CAMj2}}{\partial x_k} + q_{CAMj3} \frac{\partial R_{CAM3}}{\partial x_k} + R_{CAM3} \frac{\partial q_{CAMj3}}{\partial x_k} + \frac{\partial q_{CAMj4}}{\partial x_k}, \\ & \text{where } x_k \in \{A_s, B_s, C_s, t_{T1}, t_{T2}, t_{T3}, \theta_A, \theta_B, \theta_C\} \text{ and } j \in \{1, 2, 3\} \end{aligned}$$

and

$$\frac{\partial \Delta_N}{\partial x_k} = \frac{N_1 \frac{\partial N_1}{\partial x_k} + N_2 \frac{\partial N_2}{\partial x_k} + N_3 \frac{\partial N_3}{\partial x_k}}{\sqrt{N_1^2 + N_2^2 + N_3^2}}$$

$$\frac{\partial W}{\partial x_k} = \frac{(A_1 - B_1) \frac{\partial N_1}{\partial x_k} + (A_2 - B_2) \frac{\partial N_2}{\partial x_k} + (A_3 - B_3) \frac{\partial N_3}{\partial x_k}}{\Delta_N^2 \sqrt{(A_1 - B_1)^2 + (A_2 - B_2)^2 + (A_3 - B_3)^2}} -$$

$$- 2 \frac{(A_1 - B_1) N_1 + (A_2 - B_2) N_2 + (A_3 - B_3) N_3}{\Delta_N^3 \sqrt{(A_1 - B_1)^2 + (A_2 - B_2)^2 + (A_3 - B_3)^2}} \frac{\partial \Delta_N}{\partial x_k}$$

for all x_k in the state vector.

The derivatives of \mathbf{V}_r are then given by:

$$\frac{\partial V_{rj}}{\partial x_k} = \frac{2W N_j}{\Delta_N^2} \frac{\partial \Delta_N}{\partial x_k} - \frac{2W}{\Delta_N} \frac{\partial N_j}{\partial x_k} + \frac{2N_j}{\Delta_N} \frac{\partial W}{\partial x_k},$$

for all x_k in the state vector and for $j \in \{1, 2, 3\}$.

The Jacobian \mathbf{J}_A can now be computed using equation (6.16), that is:

$$\mathbf{J}_A = 2W_A \begin{bmatrix} (V_{r1} - V'_{r1}) \frac{\partial V_{r1} - V'_{r1}}{\partial f_u} + (V_{r2} - V'_{r2}) \frac{\partial V_{r2} - V'_{r2}}{\partial f_u} + (V_{r3} - V'_{r3}) \frac{\partial V_{r3} - V'_{r3}}{\partial f_u} \\ (V_{r1} - V'_{r1}) \frac{\partial V_{r1} - V'_{r1}}{\partial f_v} + (V_{r2} - V'_{r2}) \frac{\partial V_{r2} - V'_{r2}}{\partial f_v} + (V_{r3} - V'_{r3}) \frac{\partial V_{r3} - V'_{r3}}{\partial f_v} \\ \dots \\ (V_{r1} - V'_{r1}) \frac{\partial V_{r1} - V'_{r1}}{\partial \theta_C} + (V_{r2} - V'_{r2}) \frac{\partial V_{r2} - V'_{r2}}{\partial \theta_C} + (V_{r3} - V'_{r3}) \frac{\partial V_{r3} - V'_{r3}}{\partial \theta_C} \end{bmatrix}$$

Appendix B: Apparent Contour Term

In this appendix we derive the explicit expressions for the third term of the cost function and its Jacobian to be included in the bundle adjustment model. We start by presenting the cost function in relation to the apparent contour.

According to equation (5.3) the cone that is back-projected by the camera \mathbf{K} through the conic \mathbf{C} has its 3×3 upper-left matrix given by $\mathbf{\Gamma}_A = \mathbf{K}^T \mathbf{C} \mathbf{K}$. Expanding this equation it yields:

$$\mathbf{\Gamma}_A = \begin{bmatrix} C_{11}f_u^2 & C_{11}f_u\nu + C_{12}f_u f_v & C_{11}f_u u_0 + C_{12}f_u v_0 + C_{13}f_u \\ C_{11}f_u\nu + C_{12}f_u f_v & C_{11}\nu^2 + 2C_{12}f_v\nu + C_{22}f_v^2 & C_{11}\nu u_0 + C_{12}(\nu v_0 + f_v u_0) + C_{13}\nu + C_{22}f_v v_0 + C_{23}f_v \\ C_{11}f_u u_0 + C_{12}f_u v_0 + C_{13}f_u & C_{11}\nu u_0 + C_{12}(\nu v_0 + f_v u_0) + C_{13}\nu + C_{22}f_v v_0 + C_{23}f_v & C_{11}u_0^2 + 2C_{12}u_0 v_0 + 2C_{13}u_0 + C_{22}v_0^2 + 2C_{23}v_0 + C_{33} \end{bmatrix} \quad (6.44)$$

According to equation (5.5), the cone that is tangent to the quadric mirror and has its vertex at the optical center of the pinhole camera has its upper-left 3×3 matrix given by $\mathbf{\Gamma}_B = q_{CAM44} \mathbf{Q}_{3CAM} - \mathbf{q}_{CAM} \mathbf{q}_{CAM}^T$, where the quadric mirror is expressed by the set of equations (6.28).

The elements of the cone $\mathbf{\Gamma}_B$ are then given by:

$$\mathbf{\Gamma}_B = \begin{bmatrix} q_{CAM44}q_{CAM11}^- - q_{CAM14}^2 & q_{CAM44}q_{CAM12}^- & q_{CAM44}q_{CAM13}^- \\ -q_{CAM14}q_{CAM24} & -q_{CAM14}q_{CAM24} & -q_{CAM14}q_{CAM34} \\ q_{CAM44}q_{CAM12}^- & q_{CAM44}q_{CAM22}^- & q_{CAM44}q_{CAM23}^- \\ -q_{CAM14}q_{CAM24} & -q_{CAM24}^2 & -q_{CAM24}q_{CAM34} \\ q_{CAM44}q_{CAM13}^- & q_{CAM44}q_{CAM23}^- & q_{CAM44}q_{CAM33}^- \\ -q_{CAM14}q_{CAM34} & -q_{CAM24}q_{CAM34} & -q_{CAM34}^2 \end{bmatrix} \quad (6.45)$$

Since the cone matrices are symmetric, they have only six unique elements with five degrees of freedom, due to scaling. $\Gamma_{\mathbf{A}}$ and $\Gamma_{\mathbf{B}}$ are projectively equivalent. Therefore to compare them we use the five independent ratios. The third term of the cost function can then be formulated as:

$$\begin{aligned}
J_C = & W_C \left(\frac{\Gamma_{A12}}{\Gamma_{A11}} - \frac{\Gamma_{B12}}{\Gamma_{B11}} \right)^2 + W_C \left(\frac{\Gamma_{A13}}{\Gamma_{A11}} - \frac{\Gamma_{B13}}{\Gamma_{B11}} \right)^2 + \\
& + W_C \left(\frac{\Gamma_{A22}}{\Gamma_{A11}} - \frac{\Gamma_{B22}}{\Gamma_{B11}} \right)^2 + W_C \left(\frac{\Gamma_{A23}}{\Gamma_{A11}} - \frac{\Gamma_{B23}}{\Gamma_{B11}} \right)^2 + \\
& + W_C \left(\frac{\Gamma_{A33}}{\Gamma_{A11}} - \frac{\Gamma_{B33}}{\Gamma_{B11}} \right)^2 \tag{6.46}
\end{aligned}$$

where W_C is the weight factor to this apparent contour error term.

The Jacobian of this new term is calculated by taking derivatives of equation (6.46) with respect to all the components of the state vector. We start by presenting the derivatives of the elements of the cone matrices.

$$\begin{aligned}
\frac{\partial \Gamma_{A11}}{\partial f_u} &= 2C_{11}f_u \\
\frac{\partial \Gamma_{A11}}{\partial x_k} &= 0, \text{ for all } x_k \in \{f_v, \nu, u_0, v_0, t_{H1}, t_{H2}, t_{H3}, \dots \\
&\quad \dots, \theta_1, \theta_2, \theta_3, A_s, B_s, C_s, t_{T1}, t_{T2}, t_{T3}, \theta_A, \theta_B, \theta_C\} \\
\frac{\partial \Gamma_{A12}}{\partial f_u} &= C_{11}\nu + C_{12}f_v \\
\frac{\partial \Gamma_{A12}}{\partial f_v} &= C_{12}f_u \\
\frac{\partial \Gamma_{A12}}{\partial \nu} &= C_{11}f_u \\
\frac{\partial \Gamma_{A12}}{\partial x_k} &= 0, \text{ for all } x_k \in \{u_0, v_0, t_{H1}, t_{H2}, t_{H3}, \theta_1, \theta_2, \theta_3, \dots \\
&\quad \dots, A_s, B_s, C_s, t_{T1}, t_{T2}, t_{T3}, \theta_A, \theta_B, \theta_C\} \\
\frac{\partial \Gamma_{A13}}{\partial f_u} &= C_{11}u_0 + C_{12}v_0 + 13 \\
\frac{\partial \Gamma_{A13}}{\partial u_0} &= C_{11}f_u
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \Gamma_{A13}}{\partial v_0} &= C_{12} f_u \\
\frac{\partial \Gamma_{A13}}{\partial x_k} &= 0, \text{ for all } x_k \in \{f_v, \nu, t_{H1}, t_{H2}, t_{H3}, \theta_1, \theta_2, \theta_3, \dots \\
&\quad \dots, A_s, B_s, C_s, t_{T1}, t_{T2}, t_{T3}, \theta_A, \theta_B, \theta_C\} \\
\frac{\partial \Gamma_{A22}}{\partial f_v} &= 2C_{22} f_v + 2C_{12} \nu \\
\frac{\partial \Gamma_{A22}}{\partial \nu} &= 2C_{11} \nu + 2C_{12} f_v \\
\frac{\partial \Gamma_{A22}}{\partial x_k} &= 0, \text{ for all } x_k \in \{f_u, u_0, v_0, t_{H1}, t_{H2}, t_{H3}, \theta_1, \theta_2, \theta_3, \dots \\
&\quad \dots, A_s, B_s, C_s, t_{T1}, t_{T2}, t_{T3}, \theta_A, \theta_B, \theta_C\} \\
\frac{\partial \Gamma_{A23}}{\partial f_v} &= C_{12} u_0 + C_{22} v_0 + C_{23} \\
\frac{\partial \Gamma_{A23}}{\partial \nu} &= C_{11} u_0 + C_{12} v_0 + C_{13} \\
\frac{\partial \Gamma_{A23}}{\partial u_0} &= C_{11} \nu + C_{12} f_v \\
\frac{\partial \Gamma_{A23}}{\partial v_0} &= C_{12} \nu + C_{22} f_v \\
\frac{\partial \Gamma_{A23}}{\partial x_k} &= 0, \text{ for all } x_k \in \{f_u, t_{H1}, t_{H2}, t_{H3}, \theta_1, \theta_2, \theta_3, \dots \\
&\quad \dots, A_s, B_s, C_s, t_{T1}, t_{T2}, t_{T3}, \theta_A, \theta_B, \theta_C\} \\
\frac{\partial \Gamma_{A33}}{\partial u_0} &= 2C_{11} u_0 + 2C_{12} v_0 + 2C_{13} \\
\frac{\partial \Gamma_{A33}}{\partial v_0} &= 2C_{12} v_0 + 2C_{22} v_0 + 2C_{23} \\
\frac{\partial \Gamma_{A33}}{\partial x_k} &= 0, \text{ for all } x_k \in \{f_u, f_v, \nu, t_{H1}, t_{H2}, t_{H3}, \theta_1, \theta_2, \theta_3, \dots \\
&\quad \dots, A_s, B_s, C_s, t_{T1}, t_{T2}, t_{T3}, \theta_A, \theta_B, \theta_C\}
\end{aligned}$$

and for the cone matrix Γ_B , it yields:

$$\begin{aligned}
\frac{\partial \Gamma_{Bij}}{\partial x_k} &= q_{CAM44} \frac{\partial q_{CAMij}}{\partial x_k} + q_{CAMij} \frac{\partial q_{CAM44}}{\partial x_k} - q_{CAMi4} \frac{\partial q_{CAMj4}}{\partial x_k} - \\
&\quad - q_{CAMj4} \frac{\partial q_{CAMi4}}{\partial x_k}
\end{aligned}$$

for all x_k in the state vector and where $i, j \in \{1, 2, 3\}$.

The derivative of the cost function term is then:

$$\begin{aligned}
\frac{\partial J_C}{\partial x_k} &= 2W_C \left(\frac{\Gamma_{A12}}{\Gamma_{A11}} - \frac{\Gamma_{B12}}{\Gamma_{B11}} \right) \left(\frac{\frac{\partial \Gamma_{A12}}{\partial x_k} \Gamma_{A11} - \frac{\partial \Gamma_{A11}}{\partial x_k} \Gamma_{A12}}{\Gamma_{A11}^2} - \frac{\frac{\partial \Gamma_{B12}}{\partial x_k} \Gamma_{B11} - \frac{\partial \Gamma_{B11}}{\partial x_k} \Gamma_{B12}}{\Gamma_{B11}^2} \right) + \\
&+ 2W_C \left(\frac{\Gamma_{A13}}{\Gamma_{A11}} - \frac{\Gamma_{B13}}{\Gamma_{B11}} \right) \left(\frac{\frac{\partial \Gamma_{A13}}{\partial x_k} \Gamma_{A11} - \frac{\partial \Gamma_{A11}}{\partial x_k} \Gamma_{A13}}{\Gamma_{A11}^2} - \frac{\frac{\partial \Gamma_{B13}}{\partial x_k} \Gamma_{B11} - \frac{\partial \Gamma_{B11}}{\partial x_k} \Gamma_{B13}}{\Gamma_{B11}^2} \right) + \\
&+ 2W_C \left(\frac{\Gamma_{A22}}{\Gamma_{A11}} - \frac{\Gamma_{B22}}{\Gamma_{B11}} \right) \left(\frac{\frac{\partial \Gamma_{A22}}{\partial x_k} \Gamma_{A11} - \frac{\partial \Gamma_{A11}}{\partial x_k} \Gamma_{A22}}{\Gamma_{A11}^2} - \frac{\frac{\partial \Gamma_{B22}}{\partial x_k} \Gamma_{B11} - \frac{\partial \Gamma_{B11}}{\partial x_k} \Gamma_{B22}}{\Gamma_{B11}^2} \right) + \\
&+ 2W_C \left(\frac{\Gamma_{A23}}{\Gamma_{A11}} - \frac{\Gamma_{B23}}{\Gamma_{B11}} \right) \left(\frac{\frac{\partial \Gamma_{A23}}{\partial x_k} \Gamma_{A11} - \frac{\partial \Gamma_{A11}}{\partial x_k} \Gamma_{A23}}{\Gamma_{A11}^2} - \frac{\frac{\partial \Gamma_{B23}}{\partial x_k} \Gamma_{B11} - \frac{\partial \Gamma_{B11}}{\partial x_k} \Gamma_{B23}}{\Gamma_{B11}^2} \right) + \\
&+ 2W_C \left(\frac{\Gamma_{A33}}{\Gamma_{A11}} - \frac{\Gamma_{B33}}{\Gamma_{B11}} \right) \left(\frac{\frac{\partial \Gamma_{A33}}{\partial x_k} \Gamma_{A11} - \frac{\partial \Gamma_{A11}}{\partial x_k} \Gamma_{A33}}{\Gamma_{A11}^2} - \frac{\frac{\partial \Gamma_{B33}}{\partial x_k} \Gamma_{B11} - \frac{\partial \Gamma_{B11}}{\partial x_k} \Gamma_{B33}}{\Gamma_{B11}^2} \right)
\end{aligned}$$

where x_k represents all the elements of the state vector and $j \in \{1, 2, 3\}$, which finishes the computation of Jacobian J_C .

Chapter 7

Method 3: Linear Solution for the Pose Estimation of Noncentral Catadioptric Systems

The estimation of the pose of a visual system in relation to the world reference frame is an important problem both in computer vision and in robotics since it is relevant for several applications, namely: motion estimation, structure from motion, robot navigation, self-localization, object recognition and head and body posture.

The classical approach to the problem of estimating the position and orientation of the vision system in relation to the world frame is the perspective n-point (PnP) problem. As mentioned in the introduction, the pose problem of perspective cameras is the most studied case. The problem was originally formulated by Fischler and Bolles [37] as the fitting of a data set to a pose transformation.

In this chapter we study the pose estimation problem in the case of noncentral catadioptric systems with quadric shaped mirrors. The approach presented is linear and assumes that the camera is calibrated in the sense of a general model. It is based on the derivation of a linear constraint in the pose transformation elements (nine elements of the rotation matrix and three elements of the translation vector). This analytical constraint is based on the result that the reflection point on the mirror surface belongs also to an analytical quadric whose coordinates are dependent on the mirror, the 3D world point and the optical center (proved in section 4.2).

Experimental tests performed with simulated data and real images proved that the pose can be estimated using our linear algorithm accurately.

This method was published in the 7th Omnidirectional Vision, Camera Networks and Non-classical Cameras Workshop in conjunction with the International Conference on Computer Vision (October 2007, Rio, Brazil) [53].

7.1 Problem Statement

In this section we present some notation conventions and define the pose problem.

Homogeneous coordinates are used instead of Cartesian. Points are expressed as $\mathbf{X} = [x_1 \ x_2 \ x_3 \ x_4]^T$ and quadrics by a 4×4 symmetric matrix \mathbf{Q} . A point \mathbf{X} belongs to a quadric \mathbf{Q} if it satisfy the equation $\mathbf{X}^T \mathbf{Q} \mathbf{X} = 0$.

Consider a pinhole camera whose optical center is the point \mathbf{C} and the intrinsic parameters matrix is the matrix \mathbf{K} . The mirror surface is given by the quadric \mathbf{Q} in canonical form (expressed by \mathbf{Q}_{CAM} in camera coordinates) which is positioned freely with relation to the camera. The 3D world point \mathbf{P} is imaged by the camera and its reflection point over the mirror surface is the point \mathbf{R} . Figure 7.1 shows the reflection process and the notations adopted.

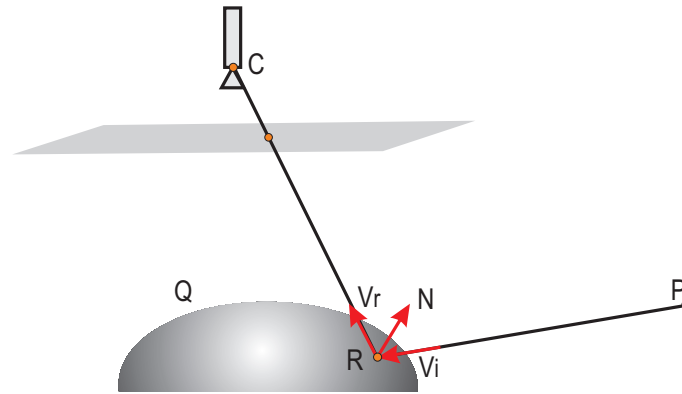


Figure 7.1: The light rays reflection and imaging in a catadioptric vision system.

The camera intrinsic parameters (focal length, principal point and skew parameter), the mirror and the pose of the camera in relation to the mirror are assumed to be known, as usually in the pose estimation problem [21, 100]. We propose to estimate the transformation \mathbf{H} between the world and camera reference frames.

7.2 Pose Estimation

In this section we develop a method to compute the pose of a calibrated catadioptric system in the world reference frame, that is to say that, the intrinsic parameters of the camera are known as well as the quadric mirror in camera coordinates. Local structure of world points is also assumed to be available.

Using back projection the reflection point \mathbf{R} is easily computed. The intrinsic parameters matrix \mathbf{K} is inverted and the reflected ray emanated

which is linear on the coordinates of the 3D point \mathbf{P} in the camera reference frame. The coefficients k_i are known since they depend on the quadric mirror coefficients, on the reflection point coordinates and on the camera optical center coordinates.

The explicit expressions of the coefficients k_i are hence given by the following equations:

$$\begin{aligned}
k_1 = & q_{CAM_{31}}c_4r_1r_2 + q_{CAM_{32}}c_4r_2^2 + q_{CAM_{33}}c_4r_2r_3 + q_{CAM_{34}}c_4r_2r_4 - \\
& - q_{CAM_{21}}c_4r_1r_3 - q_{CAM_{22}}c_4r_3r_2 - q_{CAM_{23}}c_4r_3^2 - q_{CAM_{24}}c_4r_3r_4 + \\
& + q_{CAM_{21}}c_3r_1r_4 - q_{CAM_{31}}c_2r_1r_4 + q_{CAM_{22}}c_3r_2r_4 - q_{CAM_{32}}c_2r_2r_4 + \\
& + q_{CAM_{23}}c_3r_3r_4 - q_{CAM_{33}}c_2r_3r_4 + q_{CAM_{24}}c_3r_4^2 - q_{CAM_{34}}c_2r_4^2
\end{aligned} \tag{7.2}$$

$$\begin{aligned}
k_2 = & -q_{CAM_{31}}c_4r_1^2 - q_{CAM_{32}}c_4r_1r_2 - q_{CAM_{33}}c_4r_1r_3 - q_{CAM_{34}}c_4r_1r_4 + \\
& + q_{CAM_{11}}c_4r_1r_3 + q_{CAM_{12}}c_4r_2r_3 + q_{CAM_{13}}c_4r_3^2 + q_{CAM_{14}}c_4r_3r_4 - \\
& - q_{CAM_{11}}c_3r_1r_4 + q_{CAM_{31}}c_1r_1r_4 - q_{CAM_{12}}c_3r_2r_4 + q_{CAM_{32}}c_1r_2r_4 - \\
& - q_{CAM_{13}}c_3r_3r_4 + q_{CAM_{33}}c_1r_3r_4 - q_{CAM_{14}}c_3r_4^2 + q_{CAM_{34}}c_1r_4^2
\end{aligned} \tag{7.3}$$

$$\begin{aligned}
k_3 = & q_{CAM_{21}}c_4r_1^2 + q_{CAM_{22}}c_4r_1r_2q_{CAM_{23}}c_4r_1r_3 + q_{CAM_{24}}c_4r_1r_4 - \\
& - q_{CAM_{11}}c_4r_1r_2 - q_{CAM_{12}}c_4r_2^2 - q_{CAM_{13}}c_4r_2r_3 - q_{CAM_{14}}c_4r_2r_4 + \\
& + q_{CAM_{11}}c_2r_1r_4 - q_{CAM_{21}}c_1r_1r_4 + q_{CAM_{12}}c_2r_2r_4 - q_{CAM_{22}}c_1r_2r_4 + \\
& + q_{CAM_{13}}c_2r_3r_4 - q_{CAM_{23}}c_1r_3r_4 + q_{CAM_{14}}c_2r_4^2 - q_{CAM_{24}}c_1r_4^2
\end{aligned} \tag{7.4}$$

$$\begin{aligned}
k_4 = & -q_{CAM21}c_3r_1^2 + q_{CAM31}c_2r_1^2 - q_{CAM22}c_3r_1r_2 + q_{CAM32}c_2r_1r_2 - \\
& - q_{CAM23}c_3r_1r_3 + q_{CAM33}c_2r_1r_3 - q_{CAM24}c_3r_1r_4 + q_{CAM34}c_2r_1r_4 + \\
& + q_{CAM11}c_3r_1r_2 - q_{CAM31}c_1r_1r_2 + q_{CAM12}c_3r_2^2 - q_{CAM32}c_1r_2^2 + \\
& + q_{CAM13}c_3r_2r_3 - q_{CAM33}c_1r_2r_3 + q_{CAM14}c_3r_2r_4 - q_{CAM34}c_1r_2r_4 - \\
& - q_{CAM11}c_2r_1r_3 + q_{CAM21}c_1r_1r_3 - q_{CAM12}c_2r_2r_3 + q_{CAM22}c_1r_2r_3 - \\
& - q_{CAM13}c_2r_3^2 + q_{CAM23}c_1r_3^2 - q_{CAM14}c_2r_3r_4 + q_{CAM24}c_1r_3r_4
\end{aligned} \tag{7.5}$$

where the camera optical center is $\mathbf{C} = [c_1 \ c_2 \ c_3 \ c_4]$.

On the other hand, the transformation between the camera coordinate system and the world coordinate system is given by \mathbf{H} , formally the pose we want to estimate. It transforms the 3D point as:

$$\mathbf{P} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \\ h_{31} & h_{32} & h_{33} & h_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_{w1} \\ p_{w2} \\ p_{w3} \\ p_{w4} \end{bmatrix} \tag{7.6}$$

where $\mathbf{P}_w = [p_{w1} \ p_{w2} \ p_{w3} \ p_{w4}]^T$ is the 3D point in local coordinates. The rotation matrix is obtained by performing three rotations about the coordinate axis. These Euler angles are represented by θ_1 , θ_2 and θ_3 .

Substituting in equation (7.1) the coordinates of the 3D world point given by equation (7.6), one obtains the following linear equation in the elements of the transformation matrix:

$$\begin{aligned}
& k_1 p_{w1} h_{11} + k_1 p_{w2} h_{12} + k_1 p_{w3} h_{13} + k_1 p_{w4} h_{14} + k_2 p_{w1} h_{21} + \\
& + k_2 p_{w2} h_{22} + k_2 p_{w3} h_{23} + k_2 p_{w4} h_{24} + k_3 p_{w1} h_{31} + k_3 p_{w2} h_{32} + \\
& + k_3 p_{w3} h_{33} + k_3 p_{w4} h_{34} = -k_4 p_{w4}
\end{aligned} \tag{7.7}$$

Each image point whose local 3D coordinates are known provides a different instantiation of equation (7.7). Using as many points as possible (a minimum of 12 points are needed) to enhance the robustness to noise, an over constrained system is constructed and its least squares solution, if exists, approximates the pose of the camera in the world reference frame (it is the best fit in the sense of the mean squares error).

It is then possible to compute the least squares solution that best fits the observations. The solution is thus given by the normal equations of the form $\mathbf{x} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}$. A more robust estimator can also be used if strong noise affects the solution. The least median of squares solution is a good robust estimator.

To deal with the fact that the transformation given in equation (7.6) is affine instead of rigid, additional constraints are imposed to the problem to enforce orthogonality. There are several approaches to enforce $\mathbf{H}_{\text{rot}}^T \mathbf{H}_{\text{rot}} = \mathbf{I}$, where \mathbf{H}_{rot} is the upper left 3×3 sub matrix of \mathbf{H} and \mathbf{I} is the identity matrix with the same dimension. It was not possible, however, to impose linear restrictions. This step is thus performed after solving the linear equation system. See for instance the procrustes solution in [114, 115] or the solution of the nearest orthonormal matrix due to Horn [68].

As mentioned by Hartley [58], although in general geometrical metrics provide better results than algebraic ones, in cases where the former approaches cannot be used or if some preemptive constraints are imposed, algebraic relations can perform almost ideally in terms of noise for estimation.

Experiments showed that our algebraic approach is able to estimate the pose of the camera to the world reference frame.

7.3 Experiments

In this section we present some experiments to test the validity and the robustness of the framework presented. We organized the experimental section as follows: first we study the effect of the error on the estimates by adding noise to the inputs and measuring the errors produced in the variables estimated. Next real images are used and their corresponding camera poses in the world reference frame are estimated.

Both for simulations and experiments with real images, normalized coordinates are used to avoid numerical instabilities and ill-conditioning of matrices. Robustness is achieved by normalizing the coordinates and scaling Plücker coordinates so that their norm is unitary too.

In tests with simulated data, it is important to understand how the noise affects and degrades the pose estimate. We then performed Monte Carlo tests, repeating the estimation when random Gaussian noise with zero mean was applied to each of the inputs separately, and choosing the median value for the statistics. The energy of the input error was increased so that the standard deviation from one test to the next was multiplied by a factor of 10.

Since it is also important to understand if all and how each input variable affects the solution for different types of mirrors, we performed tests with three different mirrors: a sphere, a hyperboloid and a paraboloid. The camera used was a perspective pinhole one.

Figure 7.3 shows the results obtained for the hyperbolic mirror. The x-axis scale is linear. The standard deviation of the input Gaussian noise added to variables varied from 10^{-7} (in percentage of the true value) to 10%

of the true value (which is a considerable perturbation added to the inputs). Two sets of input variables are used in the test: the image point coordinates and the 3D structure point coordinates in the world reference frame. The statistic presented is the median value of the relative error of the Euler angles (computed from the estimated rotation matrix) and the translation elements.

From the results it can be concluded that the pose parameters are robust to noise. The error of the value estimated for the pose when Gaussian noise has a 10% standard deviation is high, as expected. However, for more realistic noise energies (less than 1%) the pose estimation is accurate. Furthermore, the solution is obtained within floating point accuracy when ground truth input data is used.

Moreover, the pose parameters are more sensitive to noise added to the image point coordinates than to the noise added to the structure data of the scene.

In experiments with real images, we used two different catadioptric vision systems: a pinhole camera with a spherical mirror and with a hyperbolic mirror. The systems are both noncentral, guaranteed by off-axis positioning of the camera in relation to the quadric mirror. Figure 7.4 shows two images taken by the systems used. In what concerns to the calibration objects, we used non planar patterns to enhance the information used in the model.

The vision systems were previously calibrated in two steps. In the first step the cameras acquired images of calibration patterns and the Camera Calibration Toolbox was run [18] to calibrate the primary optical element, that is, the pinhole perspective camera. When the perspective cameras were calibrated, the known world structure (calibration patterns) applied to the iterative pose transformation was used to minimize the image reprojection error until the non linear algorithm converged to a minimum. For that purpose the method developed and described in chapter 5 to calibrate catadioptric vision systems with quadric mirrors (also published in [54]) was used. As

a result, the correspondence between pixels in the image and directions in space was provided.

As for the estimation of the pose parameters, since the ground truth pose transformation was not available we used an alternative method to compare our results with. Chen and Chang [21] algorithm was then applied to our input data. This method uses as input data the correspondences between pixels and directions in space and the 3D points in a world reference frame. Our algorithm assumes that the camera and mirror parameters are known and consequently, the direction in space corresponding to each pixel can be calculated. The results provided by both algorithms can then be compared.

We additionally evaluated the robustness of the method by computing the poses in two different positions of the catadioptric system in relation to the world and comparing the displacement induced by the two poses with the known motion applied to the catadioptric system. We hence can compare the results of our method to ground truth motion.

The methods are run for different sets of data points and the best solution is picked up from all obtained. Since the data is not ideal, we performed a nonlinear refinement of the results in order to minimize the mean error distance from the 3D point to their corresponding incident back projected ray (see figure 7.5). This step is important for our method since it is a linear method that is strongly affected by noise as expected. On the other hand, the Chen and Chang method includes a nonlinear refinement in the method itself so this step produces few effects on it.

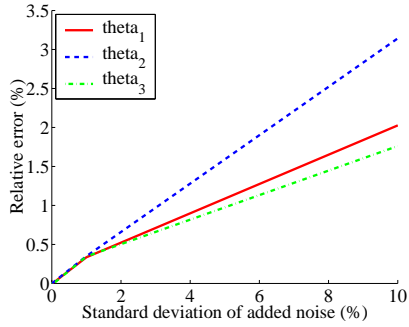
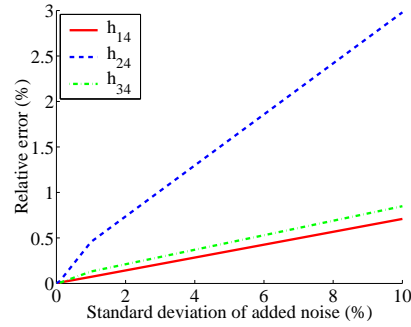
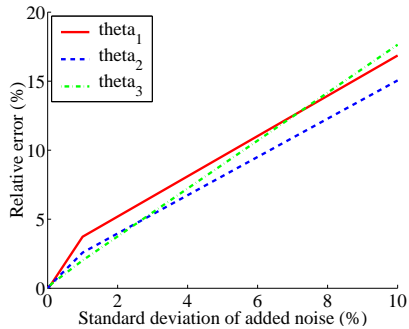
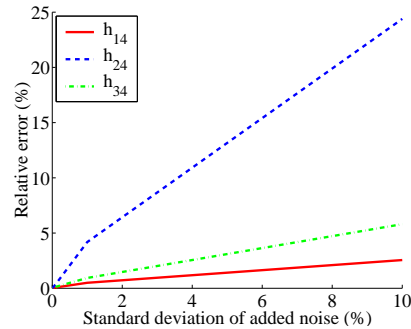
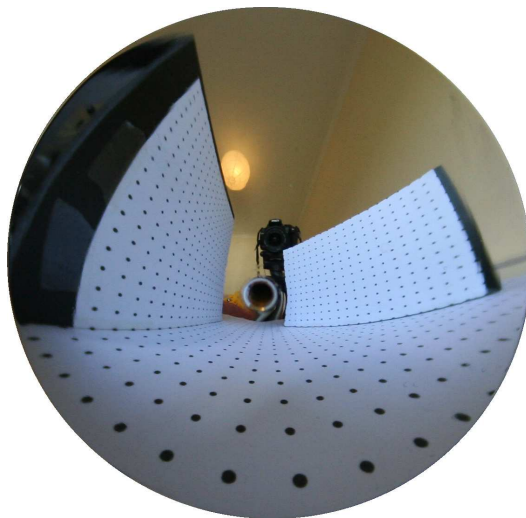
(a) θ_1 , θ_2 and θ_3 with noise added to 3D points(b) h_{14} , h_{24} and h_{34} with noise added to 3D points(c) θ_1 , θ_2 and θ_3 with noise added to image points(d) h_{14} , h_{24} and h_{34} with noise added to image points

Figure 7.3: Error analysis tests with **simulated** data using a pinhole camera and a **hyperbolic mirror**. Noise was added separately to the input data of the algorithm: image points and 3D data points. Graphics (a) and (b) plot the error for respectively the Euler angles and translation components when noise is added to the structure points and graphics (c) and (d) plot the same error measures for the case where the noise is added to the image point coordinates.



(a) Spherical mirror



(b) Hyperbolic mirror

Figure 7.4: Real images used to estimate the pose of the camera in the world reference frame.

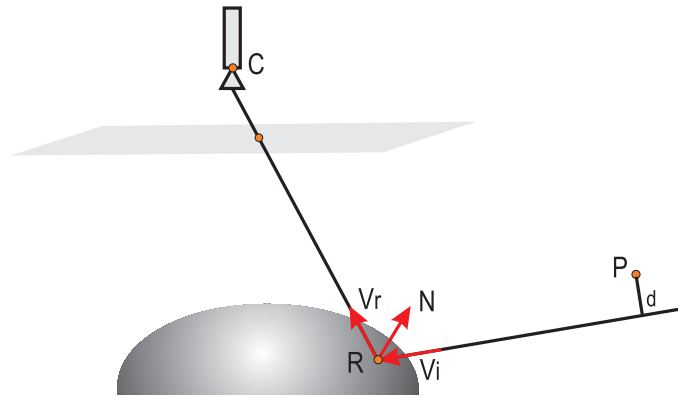


Figure 7.5: Definition of the terms in which the optimization is made for the nonlinear calibration - the distance d between a 3D point and the back projected incident ray.

The results obtained for the pose transformation (Euler rotation angles and translation elements) are presented in tables 7.1 and 7.2 and compared with the results obtained using the Chen and Chang algorithm.

We displaced the hyperbolic system by a known motion. The pose of the catadioptric system in the new position was then computed by our method and the displacement induced by the two poses (after nonlinear refinement) is compared with the known motion. The results for the estimation of the pose displacement and the ground truth motion are presented in table 7.3.

From the results it can be seen that while sensitive to noise, the linear algorithm described can estimate the pose with accuracy, presenting results

Table 7.1: Experimental tests using **real images** acquired by a pinhole camera attached to a **spherical mirror**. The pose transformation is estimated and the Euler rotation angles and translation elements are listed for the presented and Chen and Chang algorithms. The results are also compared before and after the nonlinear refinement performed to enhance accuracy. Angles are expressed in radians and displacements in *mm*.

	Before Ref.		After Ref.	
	Presented	Chen&Chang	Presented	Chen&Chang
θ_1	2.451	-0.730	-0.428	-0.714
θ_2	0.146	3.029	3.089	3.053
θ_3	0.203	-2.997	-3.131	-3.011
h_{14}	-101.138	-60.934	-57.905	-57.905
h_{24}	164.540	177.569	176.490	176.490
h_{34}	-336.094	-365.963	-366.805	-366.805

similar to those obtained with the Chen and Chang algorithm. The systems with the spherical configuration present worse results than the hyperbolic configuration and this is due to the fact that in the pre-calibration of the system the geometrical mean error obtained by the former configuration was about three times higher than the obtained with the latest one (2.0 and 0.6 mm respectively in a 400mm range). The motion between two positions in the hyperbolic configuration was also estimated with good accuracy.

7.4 Summary and Conclusions

The results obtained allow us to draw several conclusions regarding the pose estimation method described.

The pose (orientation and position) of a calibrated catadioptric system

Table 7.2: Experimental tests using **real images** acquired by a pinhole camera attached to a **hyperbolic mirror**. The pose transformation is estimated and the Euler rotation angles and translation elements are listed for the presented and Chen and Chang algorithms. The results are also compared before and after the nonlinear refinement made to enhance accuracy. Angles are expressed in radians and displacements in *mm*.

	Before Ref.		After Ref.	
	Presented	Chen&Chang	Presented	Chen&Chang
θ_1	2.135	2.115	2.129	2.129
θ_2	0.054	0.042	0.047	0.047
θ_3	0.023	0.017	0.017	0.017
h_{14}	-84.809	-82.974	-83.034	-83.034
h_{24}	126.133	126.138	126.895	126.895
h_{34}	179.138	171.421	173.559	173.559

Table 7.3: Comparison between known motion and displacement given by the two poses estimated by the method described. We used the catadioptric system with the hyperbolic mirror. Angles are expressed in radians and displacements in *mm*.

	θ_1	θ_2	θ_3	h_{14}	h_{24}	h_{34}
Known motion	0.0	0.0	0.0	20.0	20.0	0.0
Computed by our method	0.002	-0.002	0.011	20.700	19.861	-0.431

in relation to the world reference frame can be accurately estimated by a linear system of equations based on constraints defined using the correspondence of pixels and lines in space (incident light rays) and the knowledge of some structure in the world (relative positions of points). The constraints are defined based on the parameters of a quadric to which the reflection point should belong (see section 4.2). These constraints are linear in the coordinates of the 3D point projected in image.

The algorithm was compared to the Chen and Chang [21] algorithm to estimate the pose and the results obtained are similar both with real images and simulated data. The main advantage of our approach is its computer efficiency due to the linear nature of the method.

It can be concluded that the linear algorithm presented in this text allows the estimation of pose with good accuracy for noncentral catadioptric systems with quadric mirrors and that very good results can be obtained if used in conjunction to a nonlinear optimization process. Its main contributions are the proof that linear algebraic methods can be applied to those extremely nonlinear cameras, with noncentral projection and the performance of the method due to its linear nature.

Part III

DISCUSSION

Chapter 8

Discussion and Conclusions

In this chapter we discuss the main conclusions drawn throughout the thesis. A general discussion on noncentral catadioptric systems, the main theme of the work, is also presented. We start by first summarizing the work described in the text and then we discuss its main conclusions.

The vision systems used and studied in our work are noncentral catadioptric systems composed of pinhole or orthographic cameras and curved mirrors whose shape is described by a non degenerate quadric (a description that includes hyperboloids, paraboloids and ellipsoids, being spheres a particular case of the last ones). The non centrality is achieved by positioning mirror and camera in an unconstrained pose in relation to each other (location and orientation). Central configurations of the vision system and those with degenerate quadric mirrors are also suitable for the applications of our methods but they are not the object of our study.

The first topic addressed is the projection model of these vision systems. As mentioned in the introduction, as far as the author knows, there is no closed-form explicit expression that maps 3D world points to image. The projection phenomenon is then explained by either the Snell's Law or the Fermat principle in implicit multiple variable nonlinear equations which

make the performance of the projection search difficult. We then proved the existence of an additional constraint to impose to the reflection point in the projection process. This constraint allowed the description of the projection model by means of an implicit nonlinear expression on a single variable and closed-form. This mapping of a general 3D world point into image became quicker to determine and the results generally present higher accuracy.

The second topic addressed is the calibration of the visual system. We developed and presented in chapter 5 a calibration method to estimate the quadric mirror parameters (allowing the identification and classification of the mirror shape without any *a priori* information), its position and orientation in relation to the camera and also the pose of the visual system in relation to the world reference frame (extrinsic parameters). The *a priori* data are the intrinsic parameters of the pinhole perspective camera and a set of point coordinates in the world reference frame (or local for simplicity, without loss of generality). Additionally, although not preemptive, we showed that the use of the apparent mirror contour can improve the accuracy of the results.

In chapter 6 we propose another calibration method to estimate the intrinsic parameters of the pinhole camera, the parameters of the quadric mirror and the position and orientation of the camera in the world reference frame and also in relation to the mirror surface. The *a priori* data needed is a set of correspondences between an image pixel and an incident direction in space. This requirement is thus the vision system calibration in the sense of a general camera model and this method can then be regarded as a refinement of the one previously presented in the sense that the results obtained by means of the method described in chapter 5 can be used as inputs of the method of chapter 6 to improve the accuracy of the calibration. To achieve the *a priori* ray calibration for this method we also reviewed a ray calibration method that provides the needed correspondences between pixels and

direction of light rays in space.

The constraints derived for the projection model are then used in a novel method for pose estimation. In chapter 7 we present an approach to estimate the pose of the catadioptric system (perspective camera and a quadric shaped mirror) in relation to the world reference frame (extrinsic parameters) by using this algebraic constraint to derive a linear equation in the pose transformation elements between the camera and the world reference frames. We hence proved that if the camera is considered to be previously calibrated, its pose in the world reference frame can be estimated using an algebraic linear method. Procrustes or any other method is used to force the estimated transformation to produce a rigid motion (guaranteeing the orthogonality of the rotation matrix).

After this summary of the work presented in the thesis, we now draw the main conclusions of the work. Although we do not go into the details, they are summarized in the summary and conclusions section of each chapter.

The full parameterization of a catadioptric system includes the intrinsic parameters of the perspective camera (or of any other type of primary optics), its pose in relation to the specular surface (mirror), the shape parameters of the mirror itself and finally, the camera pose in relation to the world reference frame (extrinsic parameters). The total number of parameters depends on the framework adopted to the rigid transformation matrices, whether rotations are expressed by Euler angles - the standard case - or by quaternions or by 9-element rotation matrices (the affine case) with additional orthogonality constraints. The shape of the mirror can also be more or less simplified, being in the standard case expressed by non-degenerate quadric shaped surfaces. The full parameterization of a general catadioptric system is hence composed by around 20 independent parameters.

Our work described above in the text concerns the calibration of the parameters of such catadioptric systems. Whether the calibration of the

system was full or partial, some methods and insights into the problem were proposed. For the purpose of calibration, we made a lot of effort investigating the image formation phenomenon too.

The main conclusion we draw in the present thesis is that full calibration of general catadioptric systems composed of quadric shaped mirrors is difficult and extremely sensitive to many noise sources. Whereas accurate calibration can be achieved, extremely high accuracy is not easily reached for those systems. We think that this is due to the type of optical projection, particularly to non centrality and its consequence of the non existence of an explicit closed-form projection model.

The fact that the incident light rays that are projected into the image do not intersect each other in a single viewpoint makes the derivation of an explicit projection model difficult. It is not possible, as far as the author is aware of, to express explicitly the image point as function of the world point projected, in the noncentral systems studied. Therefore the calibration is often performed by using the projection of points since it is difficult to establish general properties for the projection of higher dimensional features as lines, curves in space (including planar ones as polygons and conics, and non planar ones as quartics) or surfaces. The establishment of other geometric properties as epipolar geometry or image of absolute conic is also difficult and usually provides expressions that are extremely sensitive to noise either in the calibration parameters or in the measurements.

Regarding the calibration of such systems using low-level features, that is to say, points and correspondences between image points and 3D scene points, a variety of strategies were already tested. The standard and most obvious approach (at least in the author's point of view) is the back projection of the pixel images by intersecting the emanated rays from the optical center with the specular surface. The computation of the correspondent incident ray is straightforward by computing the normal vector to surface. This

example is sufficient to visualize spatially the difference between central and noncentral systems.

In the case of central ones, there is a single point, known to be the focus of the quadric surface or of any of its longitudinal conic sections. This point (easily known) can be used to rectify the direction of the incident ray computed and this can be regarded as a reset error stage. Instead of that, the locus of the viewpoints of noncentral systems is their caustic surface, difficult to estimate for general unconstrained catadioptric systems and as a consequence they are of small help in the back projection type methods. One can intuitively understand the extremely high sensitivity to small errors of the back projection process (that includes the inversion of the intrinsic parameters of a perspective camera, a specular reflection in an uncertain surface positioned in an also uncertain position and orientation in relation to the camera).

We observed in the comprehensive set of our experiments that the parameters that have more influence in the calibration accuracy are the principal point and the orientation of the camera in relation to the mirror. On the other hand, although important as error sources, the focal length and the displacement of the camera in relation to the mirror can compensate each other by cancelling out errors, mainly if one talks about the focal length and the distance to the mirror in the optical axis direction. This fact is the ambiguity induced by the mapping between different dimensional spaces (2D and 3D).

Concerning now the metric to be used to achieve the calibration, many solutions have been already proposed (by our and other works) for this purpose. Either using geometrical or algebraic properties, the cost function built to nonlinear methods or linear equations used in linear methods are of extremely high importance to the calibration, influencing the accuracy, rate of convergence and even chance of convergence. Although geometric

distances are usually preferable to algebraic ones, since they give a direct physical metric to minimize and generally a good calibration achieved by optimizing geometric errors is undoubtedly good, algebraic distance metrics are many times easier to derive and to measure (usually provided by indirect expressions that depend on measures). Whereas physical interpretations are not always available or easy to understand, if the appropriate restrictions are imposed to the parameters (and there is a wide variety of mathematical tools to do so), the convergence can be easily achieved.

We have also observed in the experiments and simulations that the cost function is usually far from being monotonic in the parameter space. This is a real obstacle to many nonlinear iterative methods and even to linear ones. Or in other words, this is why linear methods are usually important to approximate the solution that is to be refined by nonlinear ones. Effects of cancelling out errors, as those mentioned for the focal length and distance camera-mirror and many other similar effects, most common in 20 parameter space problems, tend to give the cost function a lot of local minima and often lead to situations where the global optimal point is the deepest one of an extremely narrow hole, most difficult to find. This is the reason why continuous descent optimization strategies often stuck in a local minimum that prevents global optimization. We also concluded that these methods can be positively used in conjunction with random strategies for reset purposes and with the help of some algebraic manipulations of the cost functions.

Although difficult, if much effort and attention is put in the calibration process, the accuracy achieved can be good and led to very small geometric error. So far, in our real experiments the best full calibration ever made of a general noncentral catadioptric system was achieved using a hyperboloid mirror. The mean error for the geometric shortest distance from the 3D points to the incident calibrated ray was of about $0.6mm$ in a $400mm$ range on a set of 250 points. Partial calibration subsequently performed to the

system lead to smaller error values, however the enhancement achieved was not of big relevance (since the calibration achieved in the first stage was very good).

Recently a higher level of abstraction was attained by the models of generalized cameras. Those cameras associate each image pixel to the light ray direction responsible for it, irrespective of the path undertaken by the light (which may account for several reflections and refractions and that for catadioptric vision systems is only a specular reflection and a possible lens refraction in the primary optical element). This type of vision systems can thus model and represent almost every existing camera. They are often calibrated using the same strategies used for noncentral catadioptric systems with low-level features such as points. Higher level ones generally assume continuity of the projection model equations which is contradictory to the philosophy. In the calibration of these type of cameras we also concluded that the density of features also plays a key role in its accuracy. We observed that for sparse sets of calibration points the error tends to increase very quickly. The calibration of those systems is hence of great interest to the calibration of generalized catadioptric systems and vice-versa.

As far as the estimation of the extrinsic parameters of the vision system is concerned (partial calibration), that is to say, the pose transformation (position and orientation) of the camera in the world reference frame, a high number of solutions have been proposed (including one of our own). This most interesting topic has been studied for decades for the general perspective central camera framework. Recently some solutions to the problem in non-central cameras and in general model cameras have appeared. The solution is generally obtained by studying the geometric properties of the projection of polygons or lines, by deriving some algebraic constraints or by mixing both geometric and algebraic approaches. We observed experimentally that very high accuracy in this problem can only be achieved if the previous system

calibration (intrinsic parameters and pose of the camera in relation to the mirror) is highly accurate.

Another important conclusion drawn by the direct observation of the behavior of the calibration algorithms is the fact that some additional data can dramatically enhance the quality of the estimation results. The apparent contour of the quadric mirror, for instance, or the orthogonality of rotation matrices can be used either by restricting the parameters or by reducing the dimensionality of the parameterization.

In conclusion, our work added some contributions to the field of catadioptric vision, mainly in the study of the image formation and calibration of the system parameters. Whether full or partial, the calibration regarded as the estimation of these parameters tend to be difficult due to the nature of the projection which is noncentral. Good accuracy can however be achieved.

We also emphasize the fact that in practice, beyond the scope of our study, there are applications of those vision systems mainly to controlled robotic setups but also to real quotidian applications. It can be noticed that the majority of the visual systems used in accuracy-driven applications almost always try to guarantee the central projection property even paying the price of reduced flexibility in the design and possibility to change the optical configuration. Instead of that, for applications that don't need extremely high accurate measures, the versatility of the design and the real possibility of changing the system parameters during the application (like zooming) are often considered to be more important than the centrality of the projection. That is why many noncentral catadioptric vision systems are used is those applications.

Bibliography

- [1] Daniel Aliaga. Accurate catadioptric calibration for real-time pose estimation in room-size environments. In *IEEE International Conference on Computer Vision*, 2001.
- [2] Daniel Aliaga and Ingrid Carlbom. Fiducial planning for error-bounded pose estimation of a panoramic camera in large environments. *IEEE Robotics and Automation Magazine: Panoramic Robotics Special Issue*, June 2003.
- [3] Marc-Andre Ameller, Adrien Bartoli, and Long Quan. Reconstruction metrique minimale a partir de trois cameras affines. In *Congress AFCET de Reconnaissance des Formes et Intelligence Artificielle*, pages 471–477, January, 2002.
- [4] Adnan Ansar and Kostas Daniilidis. Linear pose estimation from points or lines. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(4), April 2003.
- [5] Helder Araújo, Rodrigo Carceroni, and Christopher Brown. A fully projective formulation to improve the accuracy of lowe pose-estimation algorithm. *Computer Vision and Image Understanding*, 70(2):227–238, 1998.

- [6] Simon Baker and Shree Nayar. A theory of single-viewpoint catadioptric image formation. *International Journal of Computer Vision*, 35(2), November 1999.
- [7] Simon Baker and Shree Nayar. *Panoramic Imaging*, chapter Single Viewpoint Catadioptric Cameras. Springer Verlag, 2001.
- [8] Hynek Bakstein and Tomas Pajdla. Non-central cameras: A review. Technical report, Czech Technical University in Prague, Prague, 2000.
- [9] João Barreto. *Imaging Beyond the Pinhole Camera*, chapter Unifying Image Plane Liftings for Central Catadioptric and Dioptric Cameras. Kluwer Academic Publishers, 2006.
- [10] João Barreto. A unifying geometric representation for central projection systems. *Computer Vision and Image Understanding*, 103(3):207–217, September 2006.
- [11] João Barreto and Helder Araújo. Issues on the geometry of central catadioptric imaging. In *IEEE Intl. Conf. on Computer Vision and Pattern Recognition*, 2001.
- [12] João Barreto and Helder Araújo. Geometric properties of central catadioptric line images. In *European Conference on Computer Vision*, 2002.
- [13] João Barreto and Helder Araújo. Paracatadioptric camera calibration using lines. In *International Conference on Computer Vision*, Nice, France, 2003.
- [14] João Barreto and Helder Araújo. A general framework for selecting world coordinate systems in perspective and catadioptric imaging applications. *International Journal of Computer Vision*, April 2004.

- [15] João Barreto and Helder Araújo. Geometric properties of central catadioptric line images and their application in calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1327–1333, August 2005.
- [16] João Barreto and Helder Araújo. Fitting conics to paracatadioptric projection of lines. *Computer Vision and Image Understanding*, 101(3):151–165, March 2006.
- [17] João Barreto and Kostas Daniilidis. Fundamental matrix for cameras with radial distortion. In *10th International Conference on Computer Vision*, Beijing, China, October 2005.
- [18] Jean-Yves Bouguet. Camera calibration toolbox for matlab. Internet web page: http://www.vision.caltech.edu/bouguetj/calib_doc/.
- [19] V. Caglioti, P. Taddei, G. Boracchi, S. Gasparin, and A. Gius. Single-image calibration of off-axis catadioptric cameras using lines. In *7th Workshop on Omnidirectional Vision*, Rio, Brazil, October 2007.
- [20] Vincenzo Caglioti and Simone Gasparini. How many planar viewing surfaces are there in noncentral catadioptric cameras?: Towards single-image localization of space lines. In *IEEE Intl. Conf. on Computer Vision and Pattern Recognition*, New York, USA, 2006.
- [21] Chu-Song Chen and Wen-Yan Chang. On pose recovery for generalized visual sensors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(7):848–861, 2004.
- [22] H.H. Chen. Pose determination from line-to-plane correspondences: Existence condition and closed-form solutions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(6):514–529, 1991.

- [23] Min Chen and James Arvo. Theory and application of specular path perturbation.
- [24] D. Weinshall D. Feldman, T. Pajdla. On the epipolar geometry of the crossed-slits projection. In *IEEE International Conference on Computer Vision*, pages 988–995, Nice, France, October 2003.
- [25] Kostas Daniilidis and Ameesh Makadia. Image processing in catadioptric planes: Spatiotemporal derivatives and optical flow computation. In *Workshop on Omnidirectional Vision*, 2002.
- [26] Philip David, Daniel DeMenthon, Ramani Duraiswami, and Hanan Samet. Softposit: Simultaneous pose and correspondence determination. *International Journal of Computer Vision*, 59(3):259–284, September 2004.
- [27] Daniel DeMenthon and Larry Davis. Exact and approximate solutions of the perspective-three-point problem. *IEEE Transaction Pattern Analysis and Machine Intelligence*, 14(11):1100–1105, 1992.
- [28] Daniel DeMenthon and Larry Davis. Model based object pose in 25 lines of code. *International Journal of Computer Vision*, 15:123–141, 1995.
- [29] Cédric Demonceaux and Pascal Vasseur. Markov random fields for catadioptric image processing. *Pattern Recognition Letters*, 27:1957–1967, 2006.
- [30] Cédric Demonceaux, Pascal Vasseur, and Claude Pégard. Robust attitude estimation with catadioptric vision. In *International Conference on Intelligent Robots and Systems*, Beijing, China, October 2006.
- [31] Aubrey K. Dunne, John Mallon, and Paul F. Whelan. Efficient generic calibration method for general cameras with single centre of projection.

- In *11th International Conference on Computer Vision*, Rio, Brazil, October 2007.
- [32] Laurent Dupont, Daniel Lazard, Sylvain Lazard, and Sylvain Petitjean. Near-optimal parameterization of the intersection of quadrics. In *Symposium on Computational Geometry (SoCG'03)*, San Diego, USA, June 2003.
- [33] Laurent Dupont, Sylvain Lazard, Sylvain Petitjean, and Daniel Lazard. *Towards the Robust Intersection of Implicit Quadrics*, chapter 5, pages 59–68. *Uncertainty in Geometric Computations*. Kluwer Academic Publishers, 2002.
- [34] J. Fabrizio, J. Tarel, and R. Benosman. Calibration of panoramic catadioptric sensors made easier. In *IEEE Workshop on Omnidirectional Vision*, June 2002.
- [35] Jonathan Fabrizio and Jean Devars. The perspective n-point problem for catadioptric sensors: An analytical approach. In *International Conference on Computer Vision and Graphics*, Warsaw, Poland, September 2004.
- [36] Mark Fiala and Anup Basu. Panoramic stereo reconstruction using non-svp optics. *Computer Vision and Image Understanding*, 98(3):363–397, June 2005.
- [37] Martin Fischler and Robert Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM: Graphics and Image Processing*, 24(6):381–395, 1981.
- [38] David Forsyth and Jean Ponce. *Computer Vision: a modern approach*. Prentice Hall, 2002.

- [39] Stephen Friedberg, Arnold Insel, and Lawrence Spence. *Linear Algebra*. Prentice Hall, New Jersey, 3rd edition, 1997.
- [40] Yasutaka Furukawa, Amit Sethi, Jean Ponce, and David Kriegman. Robust structure and motion from outlines of smooth curved surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(2):302–315, February 2006.
- [41] Xiao-Shan Gao, Xiao-Rong Hou, Jianliang Tang, and Hang-Fei Cheng. Complete solution classification for the perspective-three-point problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(8):930–943, 2003.
- [42] Christian Gebken, Antti Tolvanen, and Gerald Sommer. Pose estimation from uncertain omnidirectional image data using line-plane correspondences. *Pattern Recognition*, 4174:587–596, 2006.
- [43] Christoph Geyer and Kostas Daniilidis. A unifying theory for central panoramic and practical implications. In *IEEE European Conference on Computer Vision*, pages 445–461, Dublin, 2000.
- [44] Christoph Geyer and Kostas Daniilidis. Structure and motion from uncalibrated catadioptric views. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 279–286, Hawaii, December 2001.
- [45] Christoph Geyer and Kostas Daniilidis. Properties of the catadioptric fundamental matrix. In *IEEE European Conference on Computer Vision*, pages 140–154, 2002.
- [46] Christoph Geyer and Kostas Daniilidis. Geometric models of rolling-shutter cameras. In *6th Workshop on Omnidirectional Vision*, Beijing, China, October 2005.

- [47] Christopher Geyer and Kostas Daniilidis. Catadioptric projective geometry. *International Journal of Computer Vision*, 45(3):223–243, 2001.
- [48] Christopher Geyer and Kostas Daniilidis. Paracatadioptric camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(4), April 2002.
- [49] Joshua Gluckman and Shree Nayar. Rectified catadioptric stereo sensors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(2):224–236, February 2002.
- [50] Nuno Gonçalves and Helder Araújo. Projection model, 3d reconstruction and rigid motion estimation from non-central catadioptric cameras. In *IEEE - 3D Data Processing, Visualization and Transmission Conference*, Thessaloniki, Greece, September 2004.
- [51] Nuno Gonçalves and Helder Araújo. Rigid motion estimation for non-central catadioptric images. In *IEEE - 17th International Conference on Pattern Recognition*, Cambridge, UK, August 2004.
- [52] Nuno Gonçalves and Helder Araújo. Estimating parameters of non-central catadioptric systems using bundle adjustment. In *6th Workshop on Omnidirectional Vision*, Beijing, China, October 2005.
- [53] Nuno Gonçalves and Helder Araújo. Linear solution for the pose estimation of noncentral catadioptric systems. In *7th Workshop on Omnidirectional Vision*, Rio, Brazil, October 2007.
- [54] Nuno Gonçalves and Helder Araújo. Low-cost method for the estimation of the shape of quadric mirrors and calibration of catadioptric cameras. *Optical Engineering*, 46(7), 2007.

- [55] Michael Grossberg and Shree Nayar. A general imaging model and a method for finding its parameters. In *IEEE International Conference on Computer Vision*, Vancouver, Canada, July 2001.
- [56] A. Gruen and T. Huang. *Calibration and Orientation of Cameras in Computer Vision*. Springer Verlag, 2001.
- [57] D. Gupta and K. Daniilidis. Planar motion of a parabolic catadioptric camera. In *IEEE Conference on Computer Vision and Pattern Recognition*, Hawaii, December 2004.
- [58] Richard Hartley. Minimizing algebraic error in geometric estimation problems. In *International Conference on Computer Vision*, Bombay, 1998.
- [59] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2000.
- [60] Eugene Hecht. *Optics*. Addison-Wesley, 1987.
- [61] R. Andrew Hicks. Designing a mirror to realize a given projection. *Journal of the Optical Society of America - JOSA A*, 22(2):323–330, 2005.
- [62] R. Andrew Hicks, Marc Millstone, and Kostas Daniilidis. Realizing any central projection with a mirror pair. *Applied Optics*, 45(28):7205–7210, 2006.
- [63] Michael Hohmeyer and Seth Teller. Determining the lines through four lines. *Journal of Graphics Tools*, 4(3):11–22, 1999.
- [64] J. Hong. Image based homing. In *International Conference on Robotics and Automation - ICRA*, May 1991.

- [65] Wei Hong, Allen Yang, Kun Huang, and Yi Ma. On symmetry and multiple-view geometry: Structure, pose and calibration from a single image. *International Journal of Computer Vision*, 60(3):241–265, 2004.
- [66] Radu Horaud, Bernard Conio, Olivier Le Boulleux, and Bernard Lacolle. An analytic solution for the perspective 4-point problem. *Computer Vision, Graphics and Image Processing*, 47(1):33–44, 1989.
- [67] Berthold Horn. Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America*, 4(4):629–642, 1987.
- [68] Berthold Horn. Closed-form solution of absolute orientation using orthonormal matrices. *Journal of the Optical Society of America - A*, 5(7):1127–1135, 1988.
- [69] Z. Hu and F. Wu. A note on the number of solution of the noncoplanar p4p problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(4):550–555, 2002.
- [70] Hong Hua, Narendra Ahuja, and Chunyu Gao. Design analysis of a high-resolution panoramic camera using conventional imagers and a mirror pyramid. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(2):356–361, February 2007.
- [71] Yubin Hung, Pen-Shu Yeh, and David Harwood. Passive ranging to known planar point sets. In *IEEE International Conference on Robotics and Automation*, pages 80–85, March 1985.
- [72] Hiroshi Ishiguro, Masashi Yamamoto, and Saburo Tsuji. Omni-directional stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):257–262, February 1992.

- [73] Qiang Ji, Mauro Costa, Robert Haralick, and Linda Shapiro. An integrated linear technique for pose estimation from different geometric features. *International Journal of Pattern Recognition and Artificial Intelligence*, June 1999.
- [74] Wei Jiang, Shigeki Sugimoto, and Masatoshi Okutomi. Omnidirectional 3d reconstruction using rotating camera with mirrors. *Systems and Computers in Japan*, 38(4):1508–1520, August 2007.
- [75] S. H. Joseph. Optimal pose estimation in two and three dimensions. *Computer Vision and Image Understanding*, 72(2):215–231, February 1999.
- [76] Juho Kannala and Sami S. Brandt. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(8):1335–1340, August 2006.
- [77] Michael Kazhdan. An approximate and efficient method for optimal rotation alignment of 3d models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(7):1221–1229, July 2007.
- [78] S. Kuthirummal and S. K. Nayar. Flexible mirror imaging. In *7th Workshop on Omnidirectional Vision*, Rio, Brazil, October 2007.
- [79] Sujit Kuthirummal and Shree Nayar. Multiview radial catadioptric imaging for scene capture. *ACM Transactions on Graphics*, 25(3):916–923, 2006.
- [80] T. Lam. *Algebraic Theory of Quadratic Forms*. W.A. Benjamin, Reading, MA, 1973.

- [81] Sylvain Lazard, Luis Peñaranda, and Sylvain Petitjean. Intersecting quadrics: An efficient and exact implementation. *Computational Geometry: Theory and Applications (special issue on SoCG'04)*, 35(1-2):74–99, 2006.
- [82] Joshua Levin. A parametric algorithm for drawing pictures of solid objects composed of quadric surfaces. *Communications of the ACM*, 19(10):555–563, 1976.
- [83] Joshua Levin. Mathematical models for determining the intersection of quadric surfaces. *Computer Graphics and Image Processing*, 11(1), 1979.
- [84] Maxime Lhuillier. Automatic structure and motion using a catadioptric camera. In *6th Workshop on Omnidirectional Vision*, Beijing, China, October 2005.
- [85] Maxime Lhuillier. Automatic scene structure and camera motion using a catadioptric system. *Computer Vision and Image Understanding*, 109(2):186–203, 2008.
- [86] Shih-Schön Lin and Ruzena Bajcsy. Single-view-point omnidirectional catadioptric cone mirror imager. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(5):840–845, May 2006.
- [87] Seppo Linnainmaa, David Harwood, and Larry Davis. Pose determination of a three-dimensional object using triangle pairs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10(5):634–647, 1988.
- [88] David Lowe. Fitting parameterized three-dimensional models to images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(5):441–450, May 1991.

- [89] C. López-Franco and E. Bayro-Corrochano. Omnidirectional robot vision using conformal geometric computing. *Journal of Mathematical Imaging and Vision*, 26(3):243–260, December 2006.
- [90] Song De Ma. Conics-based stereo, motion estimation, and pose determination. *International Journal of Computer Vision*, 10(1):7–25, February 1993.
- [91] Ameesh Makadia and Kostas Daniilidis. Direct 3d-rotation estimation from spherical images via a generalized shift theorem. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2003.
- [92] Ameesh Makadia and Kostas Daniilidis. Rotation recovery from spherical images without correspondences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(7), July 2006.
- [93] Tomohiro Mashita, Yoshio Iwai, and Masahiko Yachida. Calibration method for misaligned catadioptric camera. *IEICE Transactions on Information and Systems*, 89(7):1984–1993, 2006.
- [94] Branislav Micusik and Tomas Pajdla. Autocalibration & 3d reconstruction with non-central catadioptric cameras. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 58–65, 2004.
- [95] Branislav Micusik and Tomas Pajdla. Structure from motion with wide circular field of view cameras. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(7):1135–1149, July 2006.
- [96] Francesc Moreno-Noguer, Vincent Lepetit, and Pascal Fua. Accurate non-iterative $O(n)$ solution to the pnp problem. In *11th International Conference on Computer Vision*, Rio, Brazil, October 2007.
- [97] V. Nalwa. A true omnidirectional viewer. Technical Report NJ 07733, Bell Laboratories, Holmdel, USA, 1996.

- [98] Shree Nayar and Simon Baker. Catadioptric image formation. In *DARPA Image Understanding Workshop*, New Orleans, May 1997.
- [99] Shree Nayar and Venkata Peri. *Panoramic Imaging*, chapter Folded Catadioptric Cameras. Springer Verlag, 2001.
- [100] David Nister. A minimal solution to the generalised 3-point pose problem. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2004.
- [101] David Nister and Frederik Schaffalitzky. Four points in two or three calibrated views: Theory and practice. *International Journal of Computer Vision*, 67(2):211–231, 2006.
- [102] Denis Oberkampf, Daniel DeMenthon, and Larry Davis. Iterative pose estimation using coplanar feature points. *Computer Vision and Image Understanding*, 63(3):495–511, 1996.
- [103] Tomas Pajdla. Epipolar geometry of some non-classical cameras. In *Computer Vision Winter Workshop*, pages 223–233, Slovenia, February 2001.
- [104] Tomas Pajdla. Stereo with oblique cameras. *International Journal of Computer Vision*, 2002.
- [105] Shmuel Peleg, Moshe Ben-Ezra, and Yael Pritch. Omnistere: Panoramic stereo imaging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(3):279–290, March 2001.
- [106] Robert Pless. Using many cameras as one. In *IEEE Intl. Conf. on Computer Vision and Pattern Recognition*, 2003.

- [107] Long Quan and Zhongdan Lan. Linear n-point camera pose determination. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(8):774–780, August 1999.
- [108] Bart Lamiroy Radu Horaus, Fadi Dornaika and Stéphane Christy. Object pose: The link between weak perspective, paraperspective and full perspective. *International Journal of Computer Vision*, 22(2), 1997.
- [109] Srikumar Ramalingam, Suresh Lodha, and Peter Sturm. A generic structure-from-motion framework. *Computer Vision and Image Understanding*, 103:218–228, 2006.
- [110] Srikumar Ramalingam, Peter Sturm, and S. Lodha. Towards generic self-calibration of central cameras. In *Workshop on Omnidirectional Vision*, Beijing, China, October 2005.
- [111] Srikumar Ramalingam, Peter Sturm, and Suresh Lodha. Towards complete generic camera calibration. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'05)*, 2005.
- [112] Bodo Rosenhahn, Thomas Brox, and Joachim Weickert. Three-dimensional shape knowledge for joint image segmentation and pose tracking. *Computer Vision and Image Understanding*, 73(3), 2007.
- [113] Davide Scaramuzza, Agostino Martinelli, and Roland Siegwart. A toolbox for easily calibrating omnidirectional cameras. In *IEEE Conf. Intelligent Robots and Systems*, pages 5695–5701, October 2006.
- [114] Peter Schönemann. A generalized solution of the orthogonal procrustes problem. *Psychometrika*, 31(1), March 1966.
- [115] Peter Schönemann. Fitting one matrix to another under choice of a central dilation and a rigid motion. *Psychometrika*, 35(2), June 1970.

- [116] Gerald Schweighofer and Axel Pinz. Robust pose estimation from a planar target. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(12):2024–2030, December 2006.
- [117] J. Semple and G. Kneebone. *Algebraic Projective Geometry*. Oxford University Press, London, 1959.
- [118] Ken Shoemake. Animating rotation with quaternion curves. In *Computer Graphics*, volume 19, pages 245–254. Proceedings of. SIGGRAPH 85, 1985.
- [119] Libor Spacek. A catadioptric sensor with multiple viewpoints. *Robotics and Autonomous Systems*, 51(1):3–15, April 2005.
- [120] Henrik Stewénus, David Nistér, Magnus Oskarsson, and Kalle Åström. Solutions to minimal generalized relative pose problems. In *6th Workshop on Omnidirectional Vision*, Beijing, China, 2005.
- [121] Jorge Stolfi. *Oriented Projective Geometry*. Academic Press, 1991.
- [122] Peter Sturm. Multi-view geometry for general camera models. In *IEEE Intl. Conf. on Computer Vision and Pattern Recognition*, pages 206–212, San Diego, California, June 2005.
- [123] Peter Sturm and Thomas Bonfort. How to compute the pose of an object without a direct view? In *Asian Conference on Computer Vision*, India, January 2006.
- [124] Peter Sturm and Srikumar Ramalingam. A generic camera calibration concept. In *IEEE European Conference on Computer Vision*, pages 1–13, Prague, May 2004.
- [125] Tomas Svoboda and Tomas Pajdla. Epipolar geometry for central catadioptric cameras. *International Journal of Computer Vision*, 2002.

- [126] Rahul Swaminathan, Michael Grossberg, and Shree Nayar. Caustics of catadioptric cameras. In *IEEE International Conference on Computer Vision*, July 2001.
- [127] Rahul Swaminathan, Michael Grossberg, and Shree Nayar. Designing mirrors for catadioptric systems that minimize image errors. In *5th Workshop on Omnidirectional Vision*, May 2004.
- [128] Rahul Swaminathan, Michael Grossberg, and Shree Nayar. Non-single viewpoint catadioptric cameras: Geometry and analysis. *International Journal of Computer Vision*, 66(3):211–229, 2006.
- [129] Jean-Philippe Tardif, Peter Sturm, and Sébastien Roy. Self-calibration of a general radially symmetric distortion model. In *9th European Conference on Computer Vision*, pages 186–199, Graz, Austria, May 2006.
- [130] Jean-Philippe Tarel and David Cooper. The complex representation of algebraic curves and its simple exploitation for pose estimation and invariant recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(7):663–674, July 2000.
- [131] Christel-Loic Tisse, Hugh Durrant-Whyte, and R. Andrew Hicks. An optical navigation sensor for micro aerial vehicles. *Computer Vision and Image Understanding*, 105:21–29, 2007.
- [132] Bill Triggs, Philip McLauchlan, Richard Hartley, and Andrew Fitzgibbon. Bundle adjustment - a modern synthesis. In *Vision Algorithms*, 1999.
- [133] F. Uhlig. Simultaneous block diagonalization of two real symmetric matrices. *Linear Algebra and Its Applications*, 7:281–289, 1973.

- [134] F. Uhlig. A canonical form for a pair of real symmetric matrices that generate a nonsingular pencil. *Linear Algebra and Its Applications*, 14:189–209, 1976.
- [135] John Vince. *Essential Mathematics for Computer Graphics fast*. Essential Series. Springer-Verlag, London, 2001.
- [136] Eric W. Weisstein. Point-line distance–3-dimensional. From MathWorld–A Wolfram Web Resource. <http://mathworld.wolfram.com/Point-LineDistance3-Dimensional.html>.
- [137] Y. Yagi and S. Kawato. Panoramic scene analysis with conic projection. In *International Conference on Robots and Systems - IROS*, 1990.
- [138] K. Yamazawa, Y. Yagi, and M. Yachida. Obstacle avoidance with omnidirectional image sensor hyperomni vision. In *International Conference on Robotics and Automation - ICRA*, pages 1062–1076, May 1995.
- [139] Xianghua Ying and Zhanyi Hu. Catadioptric camera calibration using geometric invariants. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(10):1260–1271, 2004.