Estimating Parameters of Non-Central Catadioptric Systems Using Bundle Adjustment

Nuno Gonçalves and Helder Araújo
Institute of Systems and Robotics - Coimbra
University of Coimbra
Polo II - Pinhal de Marrocos
3030 COIMBRA - PORTUGAL
{nunogon,helder}@isr.uc.pt

Abstract

This paper describes a new method to calibrate the intrinsic and extrinsic parameters of a generalized catadioptric camera (central or non-central) if correspondences between incident lines in space and pixels are provided (black box model calibration) in an arbitrary world reference frame. The parameters calibrated are the intrinsic parameters of the pinhole camera, the coefficients of the mirror expressed by a quadric, the position of optical center of the camera in the world reference frame and its relative orientation. A projection model relaxing Snell's Law is derived. The deviations from the Snell's Law and the image reprojection errors are minimized by a bundle adjustment method. Rotations are expressed by quaternions. Simulations and real experiments show good accuracy and robustness for this framework. A well-behaved algorithm to automatically generate the initial guess to be used in the bundle adjustment is also presented.

1. Introduction

Some recent calibration methods provide the correspondence between pixels and directions in space. They are based in the black box model introduced by Grossberg and Nayar [3] for which the calibration is no more than this list of correspondences (pixel $\longleftrightarrow$ 3D line). The vision system is therefore considered a black model and the path of light rays is unknown as well as the reflection model.

Grossberg and Nayar [3] presented a method to calibrate in this sense general vision systems with structured light patterns and Ramalingham and Sturm [7] have also formulated a method to calibrate generalized cameras. Once the calibration is performed the results can be applied to 3D reconstruction, motion analysis and several other applications as done by Pless [5], Ramalingam at al. [6] and others.

The generalized method presented by Ramalingam and Sturm in [7] uses no more than three pictures captured by the vision system in three different positions and the local geometric description of a tridimensional calibration object to calibrate the system. It provides the correspondence between each pixel and a 3D ray in space expressed in the local coordinate system of the calibration objects in one of the three positions. That means that no information about the projection itself is present in the calibration and therefore nothing is known about the intrinsic parameters and type of the camera.

Catadioptric vision systems made up of cameras and specular surfaces are specially suitable to these types of calibration methods. Indeed, in those cases, the projection model is described by a high number of intrinsic parameters which are not considered if the calibration is performed in the sense of the back box model. Furthermore, a closed form projection model exists only when the projection is central (central catadioptric cameras, see [1, 2, 10]) and no explicit projection model is known for non-central catadioptric cameras ([4]). It is then important to find out robust and accurate methods to calibrate the intrinsic parameters of catadioptric cameras with special relevance for the non-central ones.

The class of bundle adjustment methods for camera calibration requires the knowledge of the projection model that maps 3D points to the image plane such that the jacobian of the projection equations can be evaluated. Since there is no closed-form projection model for non-central catadioptric cameras, this class of methods has not been used for calibration of non-central catadioptric cameras. The complexity and the non-linearity of the projection model of central catadioptric cameras has also prevented the use of these methods for these cameras.

We are interested in the calibration of the intrinsic and extrinsic parameters of general catadioptric cameras. A previous/first calibration in the sense of back box model is assumed, that is, we assume that the correspondence pixel $\longleftrightarrow$ 3D line is already provided by using the Ramalingam and Sturm method [7] or any other suitable method.

In this paper, we propose the application of the class of bundle adjustment methods for camera calibration to general (central or not) catadioptric cameras. The explicit computation of the jacobian of the projection equations is possible due to the relaxation of the Snell's law constraint. The

non existence of the projection closed-form equations (and hence the non existence of means to provide a reflection point in the mirror surface) is circumvented by the fact that there are available correspondences between pixels and directions in space and not between pixels and points in space. The intersections between the direction rays and the mirror surface provide the reflection points.

The intersections between the directions in space and the mirror provide the reflection points. Bundle adjustment is then applied to the projection model made up with the following parameters: position of the camera in the reference frame of the first calibration (3 parameters); orientation of the camera (4 parameters, rotation quaternions are used to reduce the number of parameters and simultaneously to decrease the non-linearity of the equations); pinhole intrinsic parameters of the camera (4 parameters) and the parameters of the quadric mirror (9 parameters). The total number of parameters of the state vector is 20.

In this paper we show that bundle adjustment methods are suitable for the calibration of general catadioptric cameras and that the convergence is generally achieved in both simulations and real experiments. Since bundle adjustment methods require an initial guess for the state vector, we also provide an automatic algorithm to compute the initial guess.

In the next section the problem is described and discussed. Some considerations regarding the mathematical tools used are presented and the notation is also presented. Next, section 3 presents the projection model relating, in closed-form, the 3D lines in space with a point in the image plane. In section 4 we discuss the minimization of the cost function by a bundle adjustment method and we also discuss and present an automatic algorithm to compute the initial guess from where the bundle adjustment minimization should start. Section 5 presents the experiments and the results obtained. Finally, section 6 contains the main conclusions and the future directions of the work.

2. Problem statement

Consider a catadioptric vision system made up of a pinhole camera whose intrinsic parameters are given by matrix K and a specular mirror surface given by the quadric $Q = \{q_{ij}\}$ with $i, j \in \{1..4\}$. The camera is positioned in point P_{CAM} expressed in the world reference frame. The orientation of the pinhole camera, is given by the unitary rotation quaternion q_{rot}. Consider now a set of correspondences between pixels (u, v) in the image plane and 3D directions expressed in the world coordinate system. The directions in space are represented here by two points A and B also expressed in the world coordinate system. Figure 1 shows the relative position of the camera and the set of directions in space representing the incident light rays.

In this paper we will use quaternions to represent orientation and rotation. Quaternions are quadruples of real numbers defined as $q = [s, (q_x, q_y, q_z)] = [s, \mathbf{v}]$. For details regarding the algebraic properties of quaternions see, for instance [9]. Quaternions are used to represent position vectors in 3D space such that a point $P = [X, Y, Z]^T$ is represented by the quaternion $P_q = [0, (X, Y, Z)]$. Quaternions can also be used to perform rotations about an axis. Suppose you have a point P_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, OY and OZ by, respectively, θ_{pitch}, θ_{yaw} and θ_{roll}. The expressions for the unitary rotation quaternion that transform the coordinates of point P_q are the following:

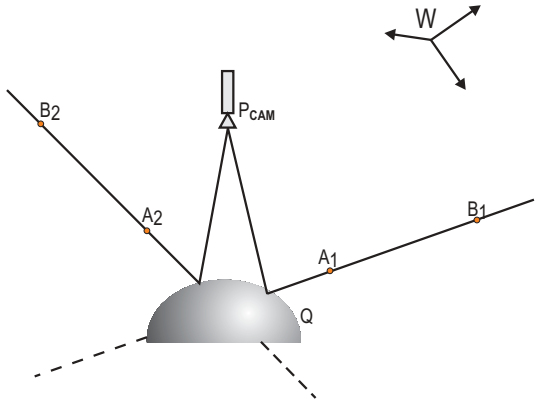

Figure 1: World reference frame and the catadioptric system.

$$s = C(\theta_y)C(\theta_p)C(\theta_r) + S(\theta_y)S(\theta_p)S(\theta_r) \quad (1)$$
$$q_x = C(\theta_y)S(\theta_p)C(\theta_r) + S(\theta_y)C(\theta_p)S(\theta_r) \quad (2)$$
$$q_y = S(\theta_y)C(\theta_p)C(\theta_r) - C(\theta_y)S(\theta_p)S(\theta_r) \quad (3)$$
$$q_z = C(\theta_y)C(\theta_p)S(\theta_r) - S(\theta_y)S(\theta_p)C(\theta_r) \quad (4)$$

where $\theta_y = \theta_{yaw}/2$, $\theta_p = \theta_{pitch}/2$ and $\theta_r = \theta_{roll}/2$.

The equation that transforms point P_q into point P'_q is:

$$P'_q = q_{rot} P_q q_{rot}^{-1} \quad (5)$$

where the multiplication of two quaternions $q_1 = [s_1, \mathbf{v_1}]$ and $q_2 = [s_2, \mathbf{v_2}]$ is given by $q_1 q_2 = \left[s_1 s_2 - \mathbf{v_1}^T \mathbf{v_2}, s_1 \mathbf{v_2} + s_2 \mathbf{v_1} + \mathbf{v_1} \times \mathbf{v_2}\right]$.

Quaternions, instead of general rotation matrices, are used since the number of parameters to estimate is fewer (4 parameters using quaternions and 9 using rotation matrices) and because, as pointed out by Triggs at al. [8] generally quaternions present better properties to apply in bundle adjustment methods since usual Euler angles present several numerical problems and pitfalls and also because quaternions have a behavior closer to linear.

In this paper we propose to calibrate the intrinsic and

extrinsic parameters of the vision system. The parameters to calibrate are:

- $K = \begin{bmatrix} f & \nu & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix}$ - intrinsic parameters of the pinhole camera
- $P_{CAM} = \begin{bmatrix} t_X & t_Y & t_Z \end{bmatrix}^T$
- $q_{rot} = [s, (q_x, q_y, q_z)]^T$
- $Q = \begin{bmatrix} q_{11} & q_{12} & q_{13} & q_{14} \\ q_{12} & q_{22} & q_{23} & q_{24} \\ q_{13} & q_{23} & q_{33} & q_{34} \\ q_{14} & q_{24} & q_{34} & 1 \end{bmatrix}$

where we consider that (1) the rotation quaternion q_{rot} is unitary, (2) the quadric surface matrix Q is symmetric $q_{ij} = q_{ji}$ and (3) since it has 9 degrees of freedom we consider the element $q_{44} = 1$.

Bundle adjustment methods require one cost function that depends on the parameters to be estimated and also requires the computation of its jacobian. Usually this class of methods has been used with projection models that map 3D points in space into image points. The cost function is often the weighted Sum of Squared Errors where the error is the algebraic 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. In the next section we derive the projection model expressions 3D line $\longleftrightarrow$ image point and its jacobian.

3. Projection model

The specular reflection is modelled by Snell's Law. According to the Snell's Law incident and reflection angles are equal. However, in our model this constraint is relaxed. Without this relaxation and restricting the projection model to the specular case, most of reflected rays would not be imaged (since they wouldn't pass at the camera optical center) and then there would not be an algebraic measure 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 quadric surface as shown in figure 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 using the angles between the reflected ray according to Snell's Law and by directly projecting it into the image (relaxing the projection law).

The quadric surface expressed by matrix Q in the world coordinate system is intersected by line ℓ defined by the 3D points A and B in cartesian coordinates. An arbitrary point

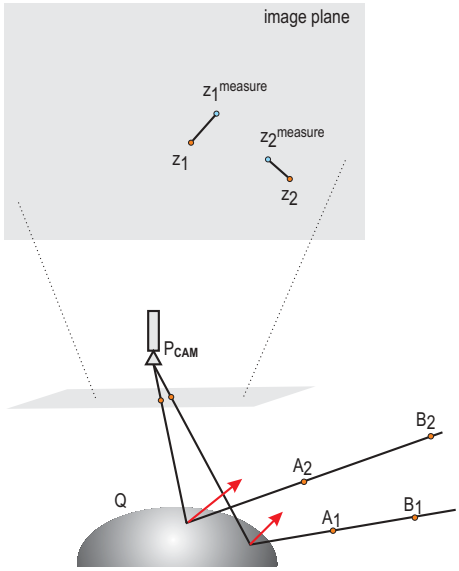

Figure 2: Projection model mapping lines in space (given by couples of points A_i and B_i) to pixels in image. Notice that Snell's law is relaxed so that the incident and reflection angles are not constrained to be equal.

in the line is expressed as $R = A + \alpha B$ where the parameter α is the solution of the incidence quadric relation:

$$\begin{bmatrix} R^T & 1 \end{bmatrix} Q \begin{bmatrix} R \\ 1 \end{bmatrix} = \begin{bmatrix} A^T + \alpha B^T & 1 \end{bmatrix} Q \begin{bmatrix} A + \alpha B \\ 1 \end{bmatrix} = 0 \tag{6}$$

Equation 6 is a quadratic equation on α that depends on the quadric mirror coefficients and also on points A and 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.

Suppose now that the unitary rotation quaternion that contains the orientation of the pinhole camera in relation to the world coordinate system is $q_{rot} = [s, (q_x, q_y, q_z)] = [s, \mathbf{v_{rot}}]$. We want to obtain the reflection point R expressed in local camera coordinates. However, since quaternions do not perform general transformations rather than pure rotations, a previous translation has to be applied to the points in the world reference frame.

The reflection point expressed in the local camera reference frame is then obtained using the equation:

$$R_{CAM_q} = q_{rot} \left((R - P_{CAM})_q \right) q_{rot}^{-1} \tag{7}$$

where $R_{CAM_q} = [0, R_{CAM}]$ is the position quaternion representing the reflection point in local camera coordinates and $(R - P_{CAM})_q$ is the position quaternion that represents

the reflection point in world coordinates translated by a vector corresponding to the position of the camera. Its expression is hence the following:

$$(R - P_{CAM})_q = \begin{bmatrix} 0, \begin{bmatrix} A_1 + \alpha B_1 - t_X \\ A_2 + \alpha B_2 - t_Y \\ A_3 + \alpha B_3 - t_Z \end{bmatrix} \end{bmatrix} \quad (8)$$

Using equation 8 in equation 7 and expanding, it yields the expression for the cartesian coordinates of the reflection point in the local camera reference frame as:

$$\begin{aligned} R_{CAM} =& (\mathbf{v_{rot}}^T (A + \alpha B - P_{CAM}))\mathbf{v_{rot}} + \\ &+ s^2 (A + \alpha B - P_{CAM}) + \\ &+ 2s\mathbf{v_{rot}} \times (A + \alpha B - P_{CAM}) \\ &- (\mathbf{v_{rot}} \times (A + \alpha B - P_{CAM})) \times \mathbf{v_{rot}} \end{aligned} \quad (9)$$

The projection of the reflection point in the image plane is given by:

$$z = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \lambda K R_{CAM} \quad (10)$$

where K is the matrix of the pinhole 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 and expanding the equations it yields:

$$\begin{cases} u = \frac{z_1}{z_3} = \frac{f R_{CAM_1} + \nu R_{CAM_2} + u_0 R_{CAM_3}}{R_{CAM_3}} \\ v = \frac{z_2}{z_3} = \frac{f R_{CAM_2} + v_0 R_{CAM_3}}{R_{CAM_3}} \end{cases} \quad (11)$$

which is the projection model for an arbitrary catadioptric system considering that the correspondences pixel $\longleftrightarrow$ incident directions 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. The errors due to the approximation in this model are minimized by the non-linear optimization algorithm, as discussed in the next section.

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 non-linear 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 [8] 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:

$$f(\mathbf{x}) = \frac{1}{2} \sum_{i=1}^{N} \{ W_I \Delta z_i(\mathbf{x})^T \Delta z_i(\mathbf{x}) + W_A \left(1 - cos\theta_i(\mathbf{x})\right)^2 \} \quad (12)$$

where $\Delta z_i(\mathbf{x}) = z_i(\mathbf{x}) - z_i^{measure}$ and $z_i(\mathbf{x}) = \begin{bmatrix} u_i & v_i \end{bmatrix}^T$ and θ_i is the angle between the reflected ray computed by the projection model of section 3 and the reflected ray computed according tp the Snell's Law. W_I and W_A are weight values applied to the reprojection error (W_I) and to the angular error (W_A). Those values can be defined arbitrarily to be the one but a better choice can be made. $\mathbf{x}$ represents the state vector whose elements are as follows:

$$\begin{aligned} \mathbf{x} = \big[& f \quad \nu \quad u_0 \quad v_0 \quad s \quad q_x \quad q_y \quad q_z \quad t_X \quad t_Y \ldots \\ & t_Z \quad q_{11} \quad q_{12} \quad q_{13} \quad q_{14} \quad q_{22} \quad q_{23} \quad q_{24} \quad q_{33} \quad q_{34} \big]^T \end{aligned} \quad (13)$$

Consider that V_r is the reflected ray computed according to the model presented in section 3 and that V_r' is the true reflected ray computed according to the Snell's Law. Assume that they are expressed in the world coordinate frame. Their equations are:

$$\begin{cases} V_r' = P_{CAM} - R = \begin{bmatrix} t_X - (a_1 + \alpha b_1) \\ t_Y - (a_2 + \alpha b_2) \\ t_Z - (a_3 + \alpha b_3) \end{bmatrix} \\ V_r = V_i - 2(V_i^T N)N \end{cases} \quad (14)$$

where V_i is the unit incident ray and N is the normal vector to the quadric surface on the reflection point R. For the incident ray we have:

$$V_i = \frac{A - B}{\|A - B\|} \quad (15)$$

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 in the reflection point R. Hence, the tangent plane is given by $\Pi_R = Q \begin{bmatrix} R & 1 \end{bmatrix}^T$ and as the direction vector of the plane Π_R is made up by the first three components of it, we have $N = Q_{3\times4} \begin{bmatrix} R & 1 \end{bmatrix}^T$, where $Q_{3\times4}$ is the rectangular matrix made up by the first three lines of the quadric mirror matrix Q.

The second element of the cost function can then by computed using the following equation for the cosine of the angle between V_r' and V_r:

$$\cos\theta_i(\mathbf{x}) = \frac{{V_r'}^T V_r}{\|V_r'\|\|V_r\|} \tag{16}$$

The computation of the derivatives of the image coordinates z_i corresponding to an arbitrary incident ray in space, with respect to the state vector components, gives us the jacobian J_1 that has to be computed. The explicit expressions for the jacobian are here omitted due to lack of space. However, they are straightforward to derive.

$$J_1 = \begin{bmatrix} \frac{\partial u}{\partial f} & \frac{\partial v}{\partial f} \\ \frac{\partial u}{\partial \nu} & \frac{\partial v}{\partial \nu} \\ \frac{\partial u}{\partial u_0} & \frac{\partial v}{\partial u_0} \\ \frac{\partial u}{\partial v_0} & \frac{\partial v}{\partial v_0} \\ \frac{\partial u}{\partial s} & \frac{\partial v}{\partial s} \\ \frac{\partial u}{\partial q_x} & \frac{\partial v}{\partial q_x} \\ \dots & \dots \\ \frac{\partial u}{\partial q_{34}} & \frac{\partial v}{\partial q_{34}} \end{bmatrix} \tag{17}$$

The jacobian of the deviation from Snell's Law J_2 is also calculated by taking the derivatives of the expression of the $cos\theta_i$ with respect to the components of the state vector. The explicit expressions are omitted due to lack of space. The cost function jacobian J is then the sum of the partial jacobians $J = J_1 + J_2$.

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 $J^T J$ is 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.

Initial Estimate

Regarding the initial guess, usually some information about the camera and the mirror is provided by the manufacturers. 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 guess of the optimal values, 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 heuristic ideas to compute one initial estimate and we emphasize that this initial guess can be enhanced whenever additional information is available.

Since the incident light directions are known at the beginning, we propose to center the quadric mirror in the point closest to all the lines in space. This point is straightforward to compute. The mirror can then be centered in this point. We propose the use of a reflecting sphere. And since no information about the dimensions of the mirror is a priori available, we can define the size of the mirror such that all the reflection points R are on one half of the surface. This provides an initial estimate for the quadric matrix in world coordinates - Q_w.

The position where to place the camera can be estimated if we compute the reflection rays using the Snell's Law applied to the incident rays on the sphere. The optical center of the camera is placed in the closest point to all the reflected rays (ideally they would intersect all in the optical center). The position of this point gives us the translation vector components (t_X, t_Y and t_Z).

The next step is to calculate the intrinsic parameters of the pinhole camera. To simplify we consider that the principal point is positioned in the center of the image (gives u_0 and v_0) and that no radial distortion exists (gives ν). For the computation of the focal length f, we consider all the reflection points R. As they are in world coordinates, and we only know the Z axis of the camera coordinate system (that coincides with the optical axis that joins the optical center and the center of the mirror) we can only compute the Z_{cam} for all the reflection points and the distance M to the optical axis. Then, since the focal length can be computed from $f = (u - u_0)Z_{cam}/X_{cam} = (v - v_0)Z_{cam}/Y_{cam}$ it can be easily proved that the focal length can by calculated by equation:

$$f = \frac{Z_{cam}}{M}\sqrt{(u - u_0)^2 + (v - v_0)^2}$$

where M is the distance between the reflection point and the optical axis. The details are omitted. The application of this equation to all available points yields different values for the focal length. We propose to use either their mean or median value. This gives f.

For the computation of the rotation quaternion we suggest to first compute the corresponding rotation matrix, and then convert it to a quaternion. The rotation matrix can be computed using a over determined system of equations since the correspondences between points in the world coordinate system (P) and in the local camera coordinate system (P') are available. If enough correspondences are available, we can estimate the rotation matrix elements inverting equation $P' = TP$, where T is the transformation matrix made up by the translation vector P_{CAM} and the rotation matrix $\{r_{ij}\}$. After estimating the rotation matrix, one can also estimate the Euler angles. Using equation 4, s, q_x, q_y and q_z can be computed.

Finally, to compute the quadric mirror coefficients $Q = q_{ij}$ expressed in local camera coordinates, we can transform the quadric by the transformation matrix already estimated: $Q = T^{-T} Q_w T^{-1}$. This gives the quadric mirror coefficients - q_{ij}. All the parameters of the state vector are now computed and an initial estimate exists.

This automatic algorithm to provide an initial guess for the bundle adjustment method can be improved if additional information about the system is available.

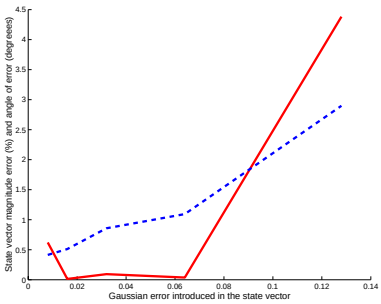

Figure 3: Results with simulation experiments. Plotted: the percent error in magnitude of the state vector (%) - solid line, and the angle between the estimated state vector and the true one, in degrees - dashed line. The gaussian white noise introduced has increasing standard deviation from 0.004 to 0.064.

In the next section experiments are presented and discussed.

5. Experiments

In this section we present experiments performed to test the validity, robustness and accuracy of the framework presented throughout this paper. Results with simulation data are first presented. They focus on the robustness of the convergence. Finally we present some results in experiments with real images using a non-central catadioptric system.

In the simulation experiments we used a catadioptric system made up of a pinhole camera with a hyperbolic mirror. Since ground truth values for the incident rays are available as well as their corresponding image points, we performed two different tests by introducing Gaussian white noise in the data.

In the first case 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. Noise was also added to the coordinates of the incident rays (obtained by the join of pairs of points A and B). However the noise added had constant energy (the variance of the noise was put at 20% of the variance of the mean point A_{mean} and B_{mean}, respectively). Figure 3 shows the relative error in the magnitude and orientation of the state vector in its vector space relative to the energy (variance) of the noise added to the input data. Table 1 shows the values obtained for the state vector in the first test, in comparison with the ground truth.

Observing the errors presented in the figure and in the table, one can conclude that in general the method converges to the true value of the state vector. Figure 4 shows the reconstructed quadric (lighter colormap) and the true one (darker colormap) for the case e_4. As can be seen in the figure, the mirror estimate is very accurate.

The second test was performed to evaluate the robustness

	True	e_1	e_2	e_3	e_4	e_5
f	800.00	798.03	799.84	803.62	800.07	775.46
ν	0.00	0.00	-0.00	0.04	0.01	0.03
u_0	320.00	322.75	315.85	320.20	314.30	328.15
v_0	240.00	240.15	243.78	238.84	246.53	253.77
s	0.98	0.98	0.98	0.98	0.98	0.98
q_x	-0.17	-0.17	-0.17	-0.17	-0.17	-0.17
q_y	0.04	0.04	0.04	0.04	0.05	0.04
q_z	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
t_x	-15.00	-15.21	-14.91	-15.01	-14.85	-16.16
t_y	0.00	0.00	0.00	0.00	-0.03	0.00
t_z	-25.00	-25.41	-25.10	-24.81	-25.04	-28.12
q_{11}	-1.06	-1.06	-1.09	-1.10	-1.02	-1.34
q_{12}	0.05	0.05	0.04	0.04	0.07	0.11
q_{13}	-0.14	-0.14	-0.14	-0.16	-0.15	-0.18
q_{14}	21.36	20.92	20.88	22.60	21.79	28.37
q_{22}	-0.85	-0.84	-0.88	-0.92	-0.81	-0.93
q_{23}	-0.61	-0.60	-0.61	-0.61	-0.63	-0.87
q_{24}	21.92	21.85	21.99	21.21	24.72	37.63
q_{33}	0.58	0.58	0.58	0.61	0.56	0.68
q_{34}	-33.15	-33.16	-33.28	-34.56	-32.22	-39.10

Table 1: Results in simulations. The results presented are relative to five different standard deviations of the gaussian white noise added to the input data - the true value of the state vector. The standard deviation of the error introduced is respectively $e_1 = 0.008$, $e_2 = 0.016$, $e_3 = 0.032$, $e_4 = 0.064$ and $e_5 = 0.128$.

and accuracy of the algorithm to generate the initial estimate of the state vector described in section 4. As the initial estimate generated by this algorithm depends only on the correspondences image point $\longleftrightarrow$ incident rays (given by pairs of points A and B), we ran the automatic initial estimate method with Gaussian white noise with several variances added to the coordinates of the points A and B. Figure 5 shows the relative error in the magnitude and orientation of the initial state vector relative to the ground truth state vector as a function of the variance of the noise added to the data. Table 2 shows the values obtained for the parameters of the system and their ground truth values. Since the error introduced is random, Monte Carlo tests were performed

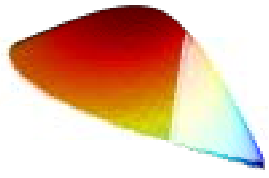

Figure 4: Estimated mirror surface (lighter colormap) and true mirror (darker colormap).

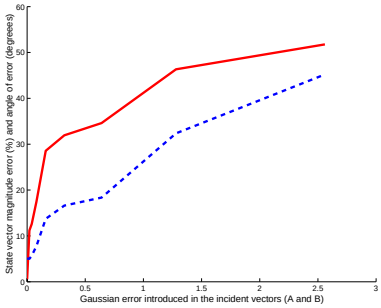

Figure 5: Relative error of the magnitude (solid line) and angle (dashed line) of the state vector estimated by the automatic initial estimate algorithm presented as function of the standard deviation of the noise added to the input data: pairs of points A and B that defines the incident direction ray. The relative error is expressed in percentage and the angle of the error is expressed in degrees.

and the results shown are the medians of the estimated values for several repetitions of the same test.

¿From the results presented in figure 5 and table 3, one can conclude that when the noise added to the input data is low, the automatic algorithm to generate the initial estimate gives very good results for the state vector. Although some of the components are first estimated with high error, the majority of the components of the state vector are estimated with good accuracy for a first estimate, noticing that no additional information is *a priori* available. As the energy of the error increases, the estimation error (either in the magnitude and in the angle) grows as expected. However, even the results obtained for 128% of standard deviation (in which case the input data is almost totally different) are acceptable.

The experiments with real images use a catadioptric camera made up by a pinhole camera and a hyperbolic mirror. The system was pre-calibrated in the sense that eighty correspondences image $\longleftrightarrow$ incident ray were available. This pre-calibration was performed using the mirror and camera values provided by manufacturers and refining them until the reprojection error becomes zero. Our aim with this initial real data is to compute the initial guess using the algorithm described in section 4 and then apply the bundle adjustment method. Figure 6 shows one of the images taken by the system.

Figure 7 shows the relative error in the magnitude and orientation of the state vector as a function of the variance of the noise added to the pre-calibration of the incident rays (assumed to be the truth values). Table 3 shows the values obtained for the parameters of the system, the initial guess and their true values.

It can be observed from the results that the error increases as the error added to the input data has increasing energy but that the method can converge to a state vector close the true one.

	True	e_1	e_2	e_3	e_4	e_5
f	800.00	798.80	803.18	638.69	492.10	418.42
ν	0.00	0.00	0.00	0.00	0.00	0.00
u_0	320.00	320.00	320.00	320.00	320.00	320.00
v_0	240.00	240.00	240.00	240.00	240.00	240.00
s	0.98	0.24	0.12	0.07	0.13	0.27
q_x	-0.17	0.10	-0.06	0.00	-0.01	0.00
q_y	0.04	-0.90	-0.96	0.93	0.41	0.85
q_z	-0.01	-0.34	-0.18	0.20	0.19	0.23
t_x	-15.00	-44.35	-13.69	-18.30	-17.07	-18.02
t_y	0.00	-8.42	18.93	16.37	18.93	15.26
t_z	-25.00	-98.33	-75.46	-78.28	-78.40	-83.18
q_{11}	-1.06	0.09	0.19	0.17	0.17	0.15
q_{12}	0.05	-0.00	0.00	0.00	-0.00	0.00
q_{13}	-0.14	-0.00	-0.00	-0.00	0.00	-0.00
q_{14}	21.36	4.10	2.61	3.16	2.77	2.80
q_{22}	-0.85	0.09	0.19	0.17	0.17	0.15
q_{23}	-0.61	-0.00	0.00	0.00	0.00	0.00
q_{24}	21.92	0.78	-3.59	-2.82	-3.14	-2.45
q_{33}	0.58	0.09	0.19	0.17	0.17	0.15
q_{34}	-33.15	9.10	14.18	13.42	13.30	12.59

Table 2: State vector values estimated by the automatic initial estimate algorithm as a function of the standard deviation of the noise added to the incident direction ray. e_i stands for the standard deviation and their values are: $e_1 = 0.02$, $e_2 = 0.04$, $e_3 = 0.16$, $e_4 = 0.32$ and $e_5 = 1.28$.

6. Conclusions

This paper presents a method to estimate the parameters of general catadioptric systems when correspondences between pixels and incident lines in space are available. A parameterized projection model relaxing the 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 directions of the incident rays.

A bundle adjustment method is applied to this model and to the data available in order to iterate 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 parameters. The computation of the initial guess and of the jacobian, both needed to the method implementation, are also addressed.

Experimental results with simulations and real images show that the method is accurate and in general converges to the global minimum. The initial guess, can be far from the optimal state vector but the bundle adjustment minimization method converges to the solution.

Future work on this method includes extensively testing

Figure 6: Real image.

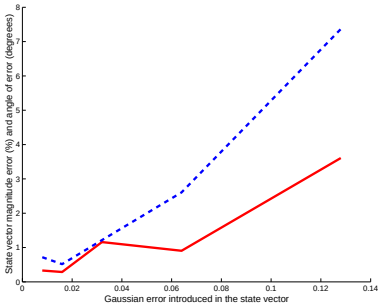

Figure 7: Real experiments. Plotted: the magnitude error of the state vector in percentage - solid line, and the angle between the state vector estimated and the true one, in degrees - dashed. The gaussian white noise introduced to the pre-calibration if the incident rays has an increasing standard deviation from $e_1 = 0.008$ to $e_5 = 0.128$.

it for several catadioptric configurations and to enhance the convergence by introducing additional constraints.

Acknowledgments

The authors gratefully acknowledge the support of projects OMNISYS–POSI/SRI/41506/2001 and CAMIDES–POSI/SRI/45970/2002, funded by the Portuguese Foundation for Science and Technology.

References

[1] João Barreto and Helder Araújo. Issues on the geometry of central catadioptric imaging. In *CVPR2001*, Haway, USA, 2001.

[2] C. Geyer and K. Daniilidis. A unifying theory for central panoramic and practical implications. In *ECCV*, pages 445–461, Dublin, 2000.

[3] Michael Grossberg and Shree Nayar. A general imaging model and a method for finding its parameters. In *ICCV01*, Vancouver, July 2001.

	True	e_1	e_2	e_3	e_4	e_5
f	800.00	777.63	793.50	842.43	726.97	614.31
ν	0.00	-0.00	0.00	0.00	0.00	0.00
u_0	320.00	317.00	320.78	322.56	332.37	278.21
v_0	240.00	240.19	239.51	243.62	228.77	238.78
s	0.98	0.98	0.98	0.98	0.98	0.98
q_x	-0.17	-0.17	-0.18	-0.18	-0.18	-0.18
q_y	0.04	0.04	0.04	0.05	0.04	0.05
q_z	-0.01	-0.01	-0.01	-0.01	-0.01	-0.01
t_x	-15.00	-15.12	-15.14	-14.73	-15.26	-14.04
t_y	0.00	0.00	-0.00	0.00	0.00	-0.00
t_z	-25.00	-25.21	-24.53	-25.24	-25.55	-29.18
q_{11}	-1.06	-1.05	-1.14	-1.07	-0.96	-1.24
q_{12}	0.05	0.05	0.06	0.06	0.05	0.06
q_{13}	-0.14	-0.15	-0.14	-0.14	-0.14	-0.15
q_{14}	21.36	21.63	21.47	22.37	23.14	21.48
q_{22}	-0.85	-0.90	-0.86	-0.87	-0.92	-0.92
q_{23}	-0.61	-0.60	-0.61	-0.62	-0.56	-0.50
q_{24}	21.92	21.89	21.08	22.13	22.00	18.45
q_{33}	0.58	0.58	0.59	0.60	0.52	0.57
q_{34}	-33.15	-33.78	-34.15	-35.48	-32.73	-42.29

Table 3: Results in real experiments. The results presented are relative to five different standard deviation of the gaussian white noise added to the input data - the true value of the state vector. The standard deviation of the error added is respectively $e_1 = 0.008$, $e_2 = 0.016$, $e_3 = 0.032$, $e_4 = 0.064$ and $e_5 = 0.128$.

[4] Branislav Micusík and Tomás Pajdla. Autocalibration and 3d reconstruction with non-central catadioptric cameras. In *CVPR 2004*, June 2004.

[5] Robert Pless. Using many cameras as one. In *CVPR03*, 2003.

[6] Srikumar Ramalingam, S. Lodha, and Peter Sturm. A generic structure-from-motion algorithm for cross-camera scenarios. In *OMNIVIS04*, 2004.

[7] Srikumar Ramalingam and Peter Sturm. A generic camera calibration concept. In *ECCV04*, pages 1–13, Prague, May 2004.

[8] Bill Triggs, Philip McLauchlan, Richard Hartley, and Andrew Fitzgibbon. *Bundle Adjustment - A Modern Synthesis*, volume LNCS 1883. Springer-Verlag, 2000.

[9] John Vince. *Essential Mathematics for Computer Graphics fast*. Essential Series. Springer-Verlag, London, 2001.

[10] Xianghua Ying and Zhanyi Hu. Catadioptric camera calibration using geometric invariants. *PAMI*, 26(10):1260–1271, October 2004.