

# Partial Calibration and Mirror Shape Recovery for Non-Central Catadioptric Systems

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

*In this paper we present a method for mirror shape recovery and partial calibration for non-central catadioptric systems. This method is based on an existing algorithm for calibration of general vision systems. In addition only two images are required for both mirror shape recovery and partial calibration instead of three as in the original algorithm. On the other hand the knowledge of the parameters of the primary vision system is required. In this paper collinearity is used to constrain the position of three points in each ray. Transformation matrices convert the local coordinates (known) into world coordinates or equivalently, estimate the motion of the calibration object between the two positions. The third point is defined as the intersection between the incident and reflected rays. In order for three points being collinear four tensor equations must be satisfied. These tensor equations depend on the motion parameters. Once the motion parameters are computed, the ray in space corresponding to each pixel can be estimated. Its intersection with the camera ray yields the coordinates of a mirror point. Simulations and real experiments showed that the solution is possible and accurate, although very sensitive to errors.*

## 1. Introduction

To overcome the limitations of the vision systems made up of perspective cameras, several different solutions have, in the recent years, been proposed and studied. Even if the usual pinhole camera is perfectly suitable for a wide variety of applications, in several cases they can be advantageously replaced by more complex optical setups involving the combination of mirrors and lenses.

The development of the new configurations for vision sensors has implied that several problems had to be tackled, namely in what concerns the development of methods and models for calibration, 3D reconstruction and others. One important problem is the development of robust, practical and easy calibration methods. On the other hand a relevant

issue is their generality.

Pinhole camera calibration methods are based on the geometric model that relates the 3D coordinates of a point with the corresponding image. The calibration is then the computation, as accurately as possible, of the projection parameters, namely the intrinsic parameters made up by the focal length, the principal point and the distortion and skew parameters, in the most common pinhole model.

Since for many of the new vision sensors their projection models are unknown, Grossberg and Nayar [2] introduced the black box camera model. For this model the correspondence between each image pixel and a 3D direction in space is estimated. The calibration of a vision system is then regarded as a list of correspondences between pixels and 3D rays in space. In [6] this calibration is assumed to be available to solve the problem of motion estimation and 3D reconstruction on a structure from motion basis and in [4] the correspondence problem is solved by assuming first a central model and then the calibrating of the system is done (the type of mirror is assumed to be known). Ramalingam and Sturm [8] presented a generic method to calibrate any vision system based on the black box model. This method uses three images of a calibration pattern to recover the motion between the coordinate frames of the camera and of the objects. Its application to omnidirectional cameras is presented in [7].

We are interested in the calibration of catadioptric vision systems. We thus present a constraint introduced to the Ramalingam and Sturm [8] method in order to reduce the amount of information used. The main change introduced to the original framework is to assume that only two images of the calibration object are available. As the method is based on collinearity constraints applied to sets of three points, and the calibration object in each position provides one point, the third point must be added. This point, if the system is considered to be a catadioptric is the reflection point on the mirror surface. Although the geometry of the mirror surface is unknown, we express this third point as a

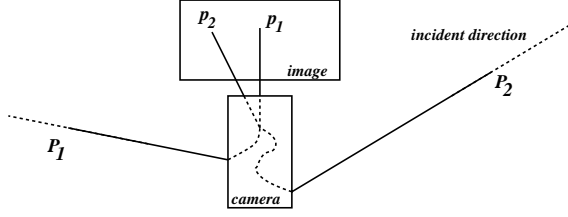


Figure 1: Black box camera with incident direction.

general point in the reflected light ray imaged by the camera and therefore its coordinates are defined up to an unknown scale factor. This method was shown to perform with good accuracy in general simulated and real experiments. Furthermore, this method provides a mean to recover the mirror surface since the scale factor is estimated. Since the method recovers sparse points on the mirror surface, the mirror shape can be then estimated using surface splines.

## 2. Related work and problem statement

The generalized camera model is used so that to each pixel corresponds one space ray as shown in figure 1 and no attention is given to the exact path taken by the light inside the vision system. This model was introduced by Groosberg and Nayar in [2].

Ramalingam and Sturm [8] introduced a calibration method for the general camera model just described. This calibration method provides the correspondence between each pixel and a ray in space. This section reviews that method.

Consider a non planar calibration object such that one knows the relative positions of the points related to each other, that is, the object model is available in a local reference frame. Three images of this object are taken by the system to be calibrated, with the object in three different positions.

Let us choose the world coordinate system so that it is centered and aligned with the coordinate frame of the object in the first position. For the pixel being considered, the calibrated 3D ray passes in one point of the calibration object in each different position. Let those points be called  $Q$ ,  $Q'$  and  $Q''$ , respectively in the first, second and third positions. Those points are in local coordinates. Figure 2 shows the coordinate systems, their relations and the calibration objects.

Consider the rotation matrix  $R'$  and the translation vector  $t'$ , as the components of the transformation matrix that rigidly transforms the local coordinate frame of the second object into the world coordinate system (that is simultaneously the first coordinate system, without loss of general-

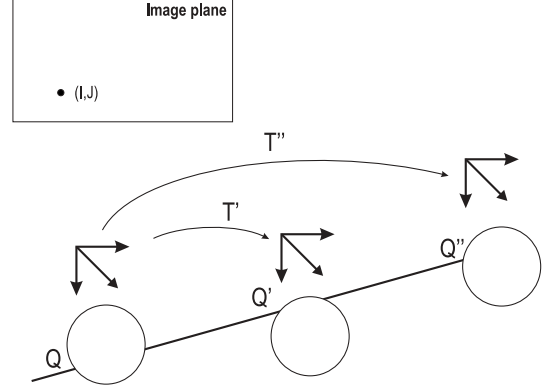


Figure 2: Visualization of the framework presented in [8].

ity). Similarly, let  $R''$  and  $t''$  be the same transformation components for the third object position (see figure 2).

Since the local points  $Q$ ,  $Q'$  and  $Q''$  coordinates are known but not the transformation matrices ( $R'$ ,  $t'$ ) and ( $R''$ ,  $t''$ ), the ray that corresponds to the pixel being considered cannot be directly computed. The method then tries to compute the transformation matrices.

The main idea of the generic calibration concept is that since all these three points are imaged in the same pixel, they must be collinear. The three points in the world coordinate system are:

$$Q \quad \begin{bmatrix} R' & t' \\ 0^T & 1 \end{bmatrix} Q' \quad \begin{bmatrix} R'' & t'' \\ 0^T & 1 \end{bmatrix} Q'' \quad (1)$$

Next consider a matrix with the expanded coordinates of the three points:

$$\begin{bmatrix} Q_1 & R'_{11}Q'_1 + R'_{12}Q'_2 + R'_{13}Q'_3 + t'_1Q'_4 & R''_{11}Q''_1 + R''_{12}Q''_2 + R''_{13}Q''_3 + t''_1Q''_4 \\ Q_2 & R'_{21}Q'_1 + R'_{22}Q'_2 + R'_{23}Q'_3 + t'_2Q'_4 & R''_{21}Q''_1 + R''_{22}Q''_2 + R''_{23}Q''_3 + t''_2Q''_4 \\ Q_3 & R'_{31}Q'_1 + R'_{32}Q'_2 + R'_{33}Q'_3 + t'_3Q'_4 & R''_{31}Q''_1 + R''_{32}Q''_2 + R''_{33}Q''_3 + t''_3Q''_4 \\ Q_4 & Q'_4 & Q''_4 \end{bmatrix} \quad (2)$$

Since the three points are collinear, as stated by [8], the rank of the matrix must be less than 3 and then all sub-determinants of size  $3 \times 3$  must be zero. This provides four equations, and each of them corresponds to a trilinear equation in point coordinates. They are also trifocal calibration tensors whose coefficients depend on the motion parameters. We call  $T^k$  to the tensor obtained by eliminating the  $k$ -th row and taking the determinant of the matrix.

The four trilinear equations in the motion parameters have known coefficients of the form  $Q_p Q'_q Q''_r$ , with  $p, q, r \in \{1..4\}$  and where 69 different unknowns can be counted. The equations written in an expanded form are presented in [9].

The four tensor equations presented can be read by the following expression:

$$\sum_1^{69} C_j T_j^k = 0 \quad (3)$$

where  $k \in \{1..4\}$ ,  $C_j$  are the 69 unknowns and  $T_j^k$  are known coefficients.

The first three tensors have 30 non-zero components and the tensor  $T^4$  has 48 non-zero components. The method presented by [8] uses only the first two tensors to recover the transformation matrices. An algorithm is presented to recover these matrices by first computing the  $C_j$  elements of tensor  $T^1$  and  $T^2$  and then by manipulating scale factors, extract from the components recovered the elements of the rotation matrices. From the rotation matrices and the same tensors, it is shown to be possible to compute the translation components. At least 29 points are used to solve the linear systems.

Since the transformation matrices between the calibration objects and the world coordinate system are recovered, the calibration step is straightforward. For each pixel, the correspondent ray in space can be obtained by the join of two calibrated points in the two of three calibration object positions. 3D reconstruction, motion analysis and other applications can be performed with this framework.

The generalized camera calibration method reviewed in this section can then be applied to several types of cameras. For central cameras or for planar or linear calibration objects some modifications to the algorithm must be made in order to be possible to recover the motion parameters and then perform the calibration itself (make the correspondence between each pixel and a ray in space). Three views of a non planar calibration object are needed for the general case.

The calibration of a generic vision system is then made using three views. In this paper we deal with catadioptric vision systems and want to calibrate them with the smallest possible number of images. As shown in the next section, the number of images required can be reduced from three to two if the primary optics is known. The problem is then how can the generic calibration method [8] can be adapted to catadioptric systems.

### 3. A Generic Catadioptric Camera Calibration

In this section we present a novel method to calibrate general catadioptric cameras that is based on the previously presented method. We show that two views of the calibration object are sufficient to calibrate those cameras.

The method in [8] uses three points of a line to define four constraints in their coordinates such that they are collinear. The collinearity of sets of three points in space then defines the conditions for the calibration of the vision system.

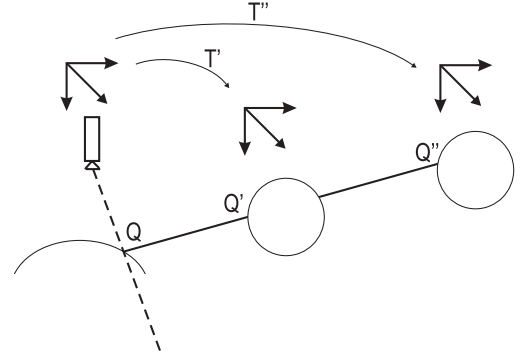


Figure 3: General visualization of the framework presented.

When using only two views of the calibration object there are only two points that are imaged in the same pixel. The collinearity constraints need however three points and thus additional information must be added.

Often the vision systems are made up by a pinhole camera and a mirror to change the light ray direction. The pinhole camera is sometimes substituted by an orthographic one. If a system of this type is considered, and if the intrinsic parameters of the camera are assumed to be known, it is possible to estimate the reflected ray that is captured by the camera. The third point needed to define the collinearity constraints belongs to this line direction.

Since the mirror geometry is unknown, to solve the problem we propose the use of a generic point in this reflected ray (see figure 3). The point considered in the reflected ray that is captured by the pinhole camera is given, after some manipulations, by the following expression:

$$Q = \begin{bmatrix} Q_1 \\ Q_2 \\ Q_3 \\ \alpha Q_4 \end{bmatrix} \quad (4)$$

where the values of  $Q_i$  can be easily computed as function of the intrinsic parameters and the image coordinates and where the unknown is now the scale factor  $\alpha$ .

Regarding the coordinates of point  $Q$  in equation 4, which is the reflection point, one can see that they differ from the coordinates of point  $Q$  used by method in [8] only in the fourth coordinate.

Similarly, the points being considered are then  $Q$  (the point where light changes its direction),  $Q'$  and  $Q''$  (the points in the incident direction that intersect the calibration object in the two positions). Multiplying these point coordinates by the corresponding transformation matrices to express all three points in the world coordinate system (notice that the point  $Q$  is already expressed in the camera and thus world coordinate system) then yields the following  $4 \times 3$  matrix:

$$\begin{bmatrix} Q_1 & R'_{11}Q'_1 + R'_{12}Q'_2 + R'_{13}Q'_3 + t'_1Q'_4 & R''_{11}Q''_1 + R''_{12}Q''_2 + R''_{13}Q''_3 + t''_1Q''_4 \\ Q_2 & R'_{21}Q'_1 + R'_{22}Q'_2 + R'_{23}Q'_3 + t'_2Q'_4 & R''_{21}Q''_1 + R''_{22}Q''_2 + R''_{23}Q''_3 + t''_2Q''_4 \\ Q_3 & R'_{31}Q'_1 + R'_{32}Q'_2 + R'_{33}Q'_3 + t'_3Q'_4 & R''_{31}Q''_1 + R''_{32}Q''_2 + R''_{33}Q''_3 + t''_3Q''_4 \\ \alpha Q_4 & Q_4 & Q_4 \end{bmatrix} \quad (5)$$

By expanding all sub-determinants and since they must be equal to zero for the points to be collinear, the four tensor expressions are similar to those in the original method, where the only difference is that where we had  $Q_4$  it is now  $\alpha Q_4$ . We do not explicitly write down those equations due to lack of space.

Since all the first three tensors  $T^1$ ,  $T^2$  and  $T^3$  depend on  $\alpha Q_4$  and since the value of  $\alpha$  is different from point to point, the number of unknowns increases dramatically. However, if one considers tensor  $T^4$ , one notices that it doesn't depend on the value of  $\alpha Q_4$  (it is the tensor obtained when the fourth row of the point matrix is suppressed).

This fact allows the computation of the solution for tensor  $T^4$  since all the coefficients are known and therefore all the unknowns of tensor  $T^4$ , from row 22 to row 69 can be linearly estimated. Since tensor  $T^4$  has 48 unknowns and they are defined up to an unknown scale factor, at least 47 points are needed to recover the tensor  $T^4$ .

The remaining unsolved three tensors still depend on the value of  $\alpha$  and our proposal is to split their equations in two parts, such that:

$$T^k = \sum_1^{69} C_j T_j^k = \sum_1^{21} C_j T_j^k + \sum_{22}^{69} C_j T_j^k = 0 \quad (6)$$

where  $k \in \{1..3\}$  and where only the second term of the right hand side depends on  $\alpha$ . Since the values of  $C_j$  for  $j \in \{22..69\}$  are now estimated by the solution of the tensor  $T^4$ , one can rewrite this equation in the form:

$$T^k = \sum_1^{69} C_j T_j^k = \sum_1^{21} C_j T_j^k + \alpha A^k = 0 \quad (7)$$

where  $A^k = \sum_{22}^{69} C_j T_j^k$ .

For each image point we then have different values of  $A^k$  and the unknown introduced is  $\alpha$ . The solution for the remaining unknowns (including the values of  $\alpha$ ), can be computed using an equation of the form of 7 for each pixel. The expressions are expanded in table 1.

Each pixel used in the estimation of the remaining coefficients provides three equations (one for each tensor) but introduces a new unknown  $\alpha_n$ . The balance is two equations for a pixel and since one has 21 unknowns (not considering the  $\alpha_n$  unknowns), at least 11 pixels are needed to recover linearly the remaining motion parameters, up to a scale factor.

The three tensors are then solved simultaneously. The extraction of the motion parameters from matrices  $T'$  and

	$C_j$	$T_j^1$	$T_j^2$	$T_j^3$
1	$R'_{11}$	0	$-Q_3 Q'_1 Q''_4$	$-Q_2 Q'_1 Q''_4$
2	$R'_{12}$	0	$-Q_3 Q'_2 Q''_4$	$-Q_2 Q'_2 Q''_4$
3	$R'_{13}$	0	$-Q_3 Q'_3 Q''_4$	$-Q_2 Q'_3 Q''_4$
4	$R'_{21}$	$-Q_3 Q'_1 Q''_4$	0	$Q_1 Q'_1 Q''_4$
5	$R'_{22}$	$-Q_3 Q'_2 Q''_4$	0	$Q_1 Q'_2 Q''_4$
6	$R'_{23}$	$-Q_3 Q'_3 Q''_4$	0	$Q_1 Q'_3 Q''_4$
7	$R'_{31}$	$Q_2 Q'_1 Q''_4$	$Q_1 Q'_1 Q''_4$	0
8	$R'_{32}$	$Q_2 Q'_2 Q''_4$	$Q_1 Q'_2 Q''_4$	0
9	$R'_{33}$	$Q_2 Q'_3 Q''_4$	$Q_1 Q'_3 Q''_4$	0
10	$R''_{11}$	0	$Q_3 Q'_4 Q''_1$	$Q_2 Q'_4 Q''_1$
11	$R''_{12}$	0	$Q_3 Q'_4 Q''_2$	$Q_2 Q'_4 Q''_2$
12	$R''_{13}$	0	$Q_3 Q'_4 Q''_3$	$Q_2 Q'_4 Q''_3$
13	$R''_{21}$	$Q_3 Q'_4 Q''_1$	0	$-Q_1 Q'_4 Q''_1$
14	$R''_{22}$	$Q_3 Q'_4 Q''_2$	0	$-Q_1 Q'_4 Q''_2$
15	$R''_{23}$	$Q_3 Q'_4 Q''_3$	0	$-Q_1 Q'_4 Q''_3$
16	$R''_{31}$	$-Q_2 Q'_4 Q''_1$	$-Q_1 Q'_4 Q''_1$	0
17	$R''_{32}$	$-Q_2 Q'_4 Q''_2$	$-Q_1 Q'_4 Q''_2$	0
18	$R''_{33}$	$-Q_2 Q'_4 Q''_3$	$-Q_1 Q'_4 Q''_3$	0
19	$t'_1 - t''_1$	0	$-Q_3 Q'_4 Q''_4$	$-Q_2 Q'_4 Q''_4$
20	$t'_2 - t''_2$	$-Q_3 Q'_4 Q''_4$	0	$Q_1 Q'_4 Q''_4$
21	$t'_3 - t''_3$	$Q_2 Q'_4 Q''_4$	$Q_1 Q'_4 Q''_4$	0
...	$\alpha_n$	$A_n^1$	$A_n^2$	$A_n^3$

Table 1: Expansion of tensors  $T^1$ ,  $T^2$  and  $T^3$  when some of the coefficients are already estimated by the tensor  $T^4$  (incorporated in the values of  $A^k$ ). Each pixel used adds a new unknown -  $\alpha_n$  and provides three equations.

$T''$  is then performed similarly to the method described in [8], including the solution for the scale factor.

Solving for the tensors  $T^1$  to  $T^3$ , provides a solution for the scale factor  $\alpha_n$  for each pixel. The fixation of the scale factor gives the reflection point on the mirror surface and therefore the algorithm provides a point on the mirror for each pixel calibrated, allowing the recovery of the mirror shape. To approximate the shape of the mirror B-cubic spline surfaces were used.

## 4. Discussion of the general solution

In this section we discuss the method just presented for solving the partial calibration and mirror shape recovery.

The existence and uniqueness of the solution depend on the type of camera and on the distribution of the calibration points in space. As noticed by [8], the calibration object should be non planar and the camera should be fully non central, to avoid the need to add additional constraints to the problem.

Since the first step of our method is to solve for tensor  $T^4$ , which has 48 unknowns, up to a scale factor, the matrix obtained to solve it should ideally have rank of 47, such that its null space is the solution required. If, however, the rank

would be 48 (due to numerical errors), the solution can be taken to be the last column of the matrix  $V$  in a singular value decomposition basis.

General simulations and real experiments suggest that full rank in the fourth tensor is generally obtained. Depending on the camera and on the calibration object (non-planar) the rank obtained for the solution of tensor  $T^4$  is generally the 47 required. The solution of the tensor  $T^4$  coefficients is then recovered up to an unknown scale factor.

For the second step, which is to solve the three tensors  $T^1$  to  $T^3$  simultaneously taking into account the solution of the first step, we also obtain generally unique solutions up to a scale factor. Both steps are thus straightforward and can be solved without the introduction of additional information.

## 5. Experiments

In this section we present the experiments performed with simulated data and real images.

We tested our method using two different catadioptric systems: a pinhole camera with a hyperbolic surface mirror and an orthographic camera with a spherical surface mirror. The former is the model of our catadioptric system used in the experiments with real images. Both these systems are non central since the camera optical center was placed away from the hyperboloid focus in the first system (in which case the system would be central, see [1, 5] for a discussion on the restrictions to impose to the relative positions of camera and mirror to achieve central projection) and since for spherical mirrors the projection is always non central.

This experimental section is divided in two parts. In the first part, using simulated data, the results obtained for the transformation matrices are presented when the full algorithm is run. Several tests were run and the results presented are representative of them. In the second part the results obtained with real images are presented.

### Simulation tests

To analyze the behavior of the estimation of the transformation matrices when data is affected by error, three sources of error were considered: the structure points of the calibration object in the second image ( $Q''$ ), the image coordinates  $(u, v)$  and the focal length  $f$ . Therefore we added gaussian white noise with zero mean in those variables, one at a time, and measured the error in the values of the estimated variables. The estimated variables estimated: the angles of rotation, the amplitude and the direction of the translation vector. Each test was repeated 20 times and the root mean square (RMS) value of the relative error was used in the plots presented.

For the errors in the rotation parameters, we decided to use the rotation angles instead of the elements of the ro-

	$(u, v)$	$f$
1	5e-3	20
2	2.5e-3	10
3	1.3e-3	5
4	6.3e-4	2.5
5	3.1e-4	1.3
6	1.6e-4	6.3e-1
7	7.8e-5	3.1e-1
8	3.9e-5	1.6e-1
9	2.0e-5	7.8e-2
10	9.8e-6	3.9e-2
11	4.9e-6	2.0e-2
12	2.4e-6	9.8e-3
13	1.2e-6	4.9e-3
14	6.1e-7	2.4e-3
15	3.1e-7	1.2e-3
16	1.5e-7	6.1e-4

Table 2: Correspondence between the number of the test and the noise added to data.

tation matrices since the number of degrees of freedom in a rotation matrix is three and their elements are related to each other. To compute their values from the rotation matrix we used a simple convergent nonlinear algorithm that approximates the given rotation matrix from three given rotation angles ([10]). Convergence was always achieved. The energy of the gaussian white noise added is relatively small since we noticed that for high values of this energy (measured by the standard deviation) the estimation errors are high.

In all the figures shown in this experimental section, the abscissa axis is logarithmic, such that the next axis value is twice the current value. The number plotted in the abscissa axis is the  $\#$  of the test. Table 2 shows the correspondence between the test number and the standard deviation of the noise added to the respective parameters.

Figure 4 plots the RMS value for the relative error in (a) the estimation of the rotation angles  $\theta_X$  and  $\theta_Z$ , (b) in the amplitude of the translation vector and (c) the RMS angle of error (in degrees) in the direction of the translation vector, when gaussian white noise with zero mean is introduced to the image coordinates  $(u, v)$  of the points in the first image. The results for the transformation matrix  $T'$  are plotted in solid lines and the results for the transformation matrix  $T''$  are plotted in dashed lines. For test  $\#k$ , the added noise has standard deviation of  $5e - 3 \times 2^{k-1}$  (see table 2 for the noise errors). The color scheme is: red -  $\theta_X$  and green -  $\theta_Z$ . For lack of space and since they are qualitatively similar to these ones, we omit the results for the estimation of the transformation matrices  $T'$  and  $T''$  when noise is added to the structure coordinates of the calibration object in the second image -  $Q''$ .

Finally, figure 5 plots the RMS value of the relative error in (a) the estimation of the rotation angles  $\theta_X$  and  $\theta_Z$ , (b) in the amplitude of the translation vector and (c) the RMS angle of error (in degrees) in the direction of the translation vector, when gaussian white noise with zero mean is added to the focal length (intrinsic parameter). The results for the transformation matrix  $T'$  are plotted in solid lines and the results for the transformation matrix  $T''$  are plotted in dashed lines. For the test  $k$ , the added noise has standard deviation of  $2e + 1 \times 2^{k-1}$  (see table 2 for the noise errors).

From the figures one can see that the estimation error rises quickly as the noise energy added to the data inputs is increased. This sensitivity is higher for the translation vector (amplitude and direction) than for the rotation angles. Furthermore, high values of estimation error are generally obtained. It is also noticed from the figures that the estimation of the transformation components is more robust to the noise in the focal length (intrinsic parameters) than to the noise in  $Q''$  and in the image coordinates  $(u, v)$ .

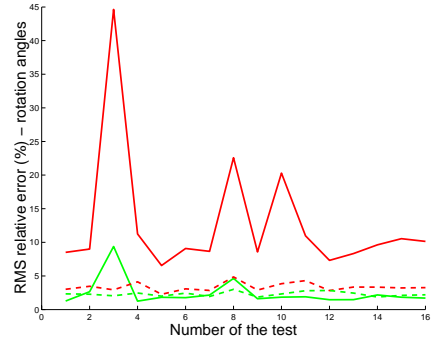
As remarked in the introduction, this framework provides a means to recover locally the shape of the mirror surface since the computation of the reflection points on the surface is possible. Point  $Q$  which is on the reflected ray is scaled by the scale factor  $\alpha$  to be on the mirror surface. By solving for tensors  $T^1, T^2$  and  $T^3$  simultaneously one also obtains the solution of the scale factor for each point (see equation in table 1). For each pixel/point used one has one point on the mirror surface. To locally recover the shape of the mirror surface several techniques can be used. For easy visualization we used cubic B-spline surfaces to fit the points set. Since we know that the mirror is hyperbolic one could fit data to an hyperboloid. However, in a realistic scenario one hasn't this information and so splines are best suitable. Figure 6 shows the reconstruction of the mirror surface as well as the true surface. It can be observed that the mirror shape is accurately recovered.

## Experiments with real images

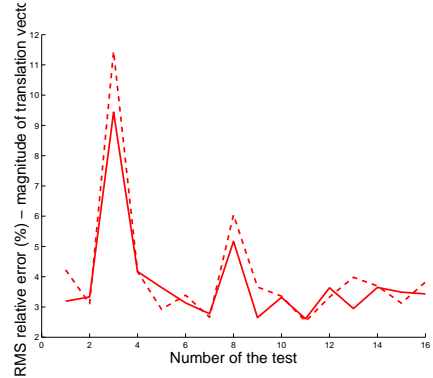
Our catadioptric system is made up of a pinhole camera and a hyperbolic mirror. Several tests were performed using those images. Figure 7 shows an image taken by the real system.

Table 3 shows the estimated values and their corresponding relative errors for two representative tests: when the motion is composed by translation in two axis (above rows) and when there are rotations in two axis (bottom rows).

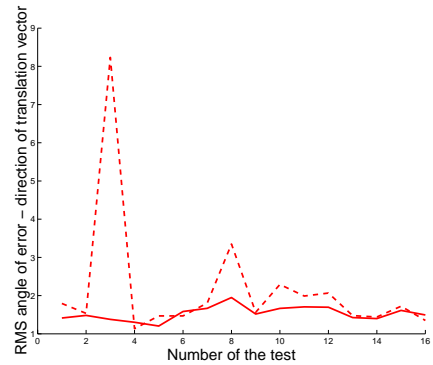
The results presented in the table 3 prove that this calibration method gives good results in the estimation of the transformation matrices that allow for the pixel-ray calibration of the catadioptric system. However, those results are somehow very sensitive to noise and better results are expected if the primary optics is more carefully calibrated.



(a) rotation angles

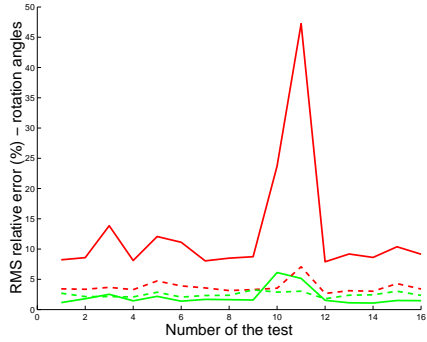


(b) amplitude of the translation vector

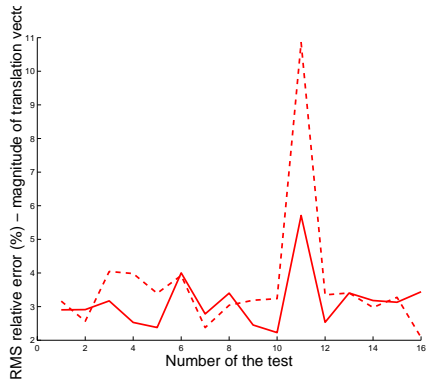


(c) direction of the translation vector

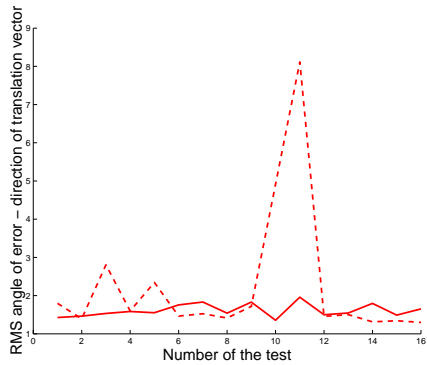
Figure 4: RMS relative error in the estimation of the transformation matrices  $T'$  and  $T''$  when gaussian white noise with zero mean is added to the image coordinates  $(u, v)$ .



(a) rotation angles



(b) amplitude of the translation vector



(c) direction of the translation vector

Figure 5: RMS relative error in the estimation of the transformation matrices  $T'$  and  $T''$  when gaussian white noise with zero mean is added to the focal length.

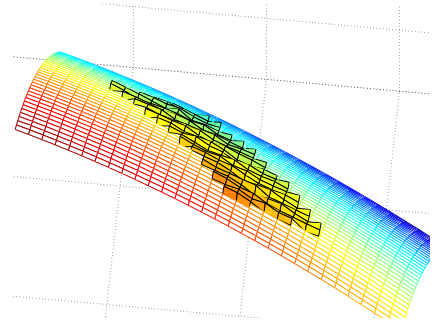


Figure 6: Local reconstruction of the mirror shape using cubic B-spline surfaces. The bigger panel is the true mirror surface and the smaller is its surface reconstructed by cubic B-splines.



Figure 7: Real image taken by our catadioptric system with a hyperbolic mirror.

## 6. Conclusion and directions

In this paper we presented modifications introduced to the generalized calibration method presented in [8]. These changes aim at taking into account restrictions that are specific to a general catadioptric vision systems. The only restriction imposed to the vision system is that it must be possible to recover the reflected light ray direction that is imaged in any pixel using no more than the intrinsic parameters of the pinhole or orthographic camera and the pixel image coordinates. These modifications rearrange the equations used in the original method in such a way that only two images are required instead of three.

Calibration is performed in the sense that for each image pixel there is a correspondent ray in space. Collinearity constraints over three points in each ray are used. If a calibration object with known local coordinates is imaged in two different positions, the points imaged in the same pixel can be used to apply the collinearity constraints. These points are known in local coordinates and two transformation matrices convert their coordinates into world coordinate (with-

GT	$\theta_x$	$\theta_y$	$\theta_z$	$t_x$	$t_y$	$t_z$
Est	-50.0	0.0	50.0	0.0	0.0	0.0
Err %	-43.59	0.09	45.64	-0.16	0.17	-0.08
	<b>12.82</b>	-	<b>8.72</b>	-	-	-
GT	0.0	0.0	0.0	$-\pi/5$	0.0	$\pi/2$
Est	2.6	-0.4	-3.9	-0.58	-0.13	1.66
Err %	-	-	-	<b>7.69</b>	-	<b>5.68</b>

Table 3: Results with real images. The rotation angle are expressed in radians and the translation coefficients in  $mm$ . GT stands for Ground truth values and Est for the estimated ones.

out loss of generality the camera coordinate system is assumed to be the world reference frame). The third point is defined by our method as belonging to the recovered ray that entered the camera. The third point is thus elsewhere in this ray. In order for those three points to be collinear they must respect four tensor equations that depend on their local coordinates and on the components of the transformation matrices, that is, they depend on the motion parameters. Once those motion parameters are recovered, the ray in space that corresponds to each pixel can thus be estimated.

The estimation process is then divided in two steps. The first step aims at recovering the tensor  $T^4$  coefficients. To recover these coefficients at least 47 points are needed. Some of the coefficients (that depend on the motion parameters) are recovered up to a scale factor.

The experiments made so far suggest that generally the first step allows the recovery of tensor  $T^4$  coefficients up to a scale factor. However, as was also noticed by [8], for central cameras or some non central configurations there are rank deficiency problems (the higher the non centrality of the camera the smaller this deficiency).

When the first step is concluded, its results are used to compute the remaining motion parameters contained in tensors  $T^1$  to  $T^3$ . All experiments performed in this second step showed that the reconstruction of those remaining parameters is stable and in general a unique solution is easily obtained, since no rank deficiency problem arises.

It was also shown that the mirror shape can be recovered using this calibration method. After recovering the transformation matrices, the coordinates of the third calibration points over the reflected ray can be recovered. As these points belong to the mirror surface, mirror shape reconstruction can be performed. We showed that with cubic spline surfaces the mirror surface can be well approximated.

The main conclusion to draw is that for non-central catadioptric systems it is possible to calibrate the system with only two images. Since the method showed to be sensitive to noise, care should be taken in the the primary optics calibration.

In the future we intend to study how the non-centrality of the system affects the estimation of the motion parameters (and thus the calibration of the system) and how additional constraints can be added to solve the problem of noise sensitivity.

## 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] Christopher Geyer and Kostas Daniilidis. Catadioptric projective geometry. *International Journal on Computer Vision*, 45(3):223–243, 2001.
- [2] Michael Grossberg and Shree Nayar. A general imaging model and a method for finding its parameters. In *ICCV01*, Vancouver, July 2001.
- [3] Mark Halstead, Brian Barsky, Stanley Klein, and Robert Mandell. Reconstructing curved surfaces from specular reflection patterns using spline surfaces fitting of normals. In *Proceedings of SIGGRAPH*, 1996.
- [4] Branislav Micusík and Tomáš Pajdla. Autocalibration and 3d reconstruction with non-central catadioptric cameras. In *CVPR 2004*, Washington US, June 2004.
- [5] Shree Nayar and Simon Baker. Catadioptric image formation. In *DARPA Image Understanding Workshop*, New Orleans, May 1997.
- [6] Robert Pless. Using many cameras as one. In *CVPR03*, 2003.
- [7] Srikumar Ramalingam, S. Lodha, and Peter Sturm. A generic structure-from-motion algorithm for cross-camera scenarios. In *OMNIVIS04*, pages 175–186, Prague, May 2004.
- [8] Srikumar Ramalingam and Peter Sturm. A generic camera calibration concept. In *ECCV04*, pages 1–13, Prague, May 2004.
- [9] Peter Sturm and Srikumar Ramalingam. A generic calibration concept - theory and algorithms. Technical Report 5058, INRIA, December 2003.
- [10] Zhengyou Zhang. A flexible new technique for camera calibration. Technical Report MSR-TR-98-71, Microsoft Research, 1998. Last time updated: 2002.