Linear solution for the pose estimation of noncentral catadioptric systems

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

Abstract

This paper presents a linear method to estimate the pose of a noncentral catadioptric system with a quadric shaped mirror in relation to a world reference frame (or local reference frame without loss of generality). The vision system is assumed to be calibrated. The method uses also as input data the structure of the scene. It is proved that any reflection point should belong to an analytical quadric that intersects the mirror quadric itself. This constraint can be written linearly in the 3D scene point coordinates (in the camera reference frame). The unknown pose screw transformation that relates camera and world reference frames can then be used in the linear model, allowing for the construction of a linear equation in the pose transformation elements. Additional constraints are used to force the estimated rotation elements to build an orthogonal matrix. Tests with simulated data and also on real images with different mirrors proved the method to be consistent and to estimate the pose accurately. However, it was also observed that the method is sensitive to noise. The results are compared with another method.

1. Introduction

Systems that use mirrors and cameras are called catadioptric and recently they have been studied and used in several applications. Amongst all catadioptric systems, those which use rotationally symmetric mirrors and in particular those whose mirrors are quadrics are probably the most used.

Catadioptric vision systems can be divided into two types depending on whether the projection is central or not, that is, depending on whether all incident light rays intersect each other in a unique viewpoint or not. In the general case, the optical center of the camera is the effective viewpoint to guarantee central projection. Particularly, it has been shown by Nayar and Baker [1] 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. However, for the general case and when this constraint is relaxed, the projection is noncentral which implies that the light rays do not intersect at an effective single viewpoint.

Noncentral vision systems do not have, in general, a projection model. As a result closed form expressions relating 3D world points coordinates to their corresponding image coordinates do not exist. However, in the case of central systems, closed form expressions exist, since their effective viewpoint is a single point.

To overcome the nonexistence of a projection model for some types of cameras, a new model of cameras has been proposed, namely generalized cameras [12, 17, 21]. This class of cameras, also called black-box cameras, associate each pixel to a direction in space. Hence, all cameras (central or not) can be modeled by a general model. Calibration of those vision systems results in a list of correspondences between a line in space and each pixel.

We are interested in catadioptric vision systems, composed by a central pinhole or orthographic camera and a curved mirror surface for panoramic images. Particularly, we are interested in those catadioptric systems with quadric shape mirrors (includes spheres, paraboloids, ellipsoids and hyperboloids). In the general case, those systems have a noncentral projection.

The estimation of the pose of a visual system relative 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 relative to the world frame is the perspective n-point (PnP) problem. The pose problem of perspective cameras is the most studied case. The problem was originally formulated by Fischler and Bolles [8] as the fitting of a data set to a pose transformation. Several solutions have been presented, depending on the number of points (or lines) that are used [9, 14, 22]. For the case of central catadioptric cameras, the pose problem has been also extensively studied by [3, 23].

For non perspective and other cameras, there are recent works either in general cameras and in more specific non-central catadioptric systems. Chen and Chang [6] 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). The solution is obtained by solving a univariate polynomial of degree 8 for the general case. If the camera is central, the solution is obtained by finding the roots of a polynomial of degree 4. On the other hand, Nister and Stewenius et al. [16, 20] presented a solution for the pose estimation of a calibrated camera by equivalently solving a polynomial of eight order. The solution is obtained by formulating the intersection of a circle and a ruled quadric surface (in an algebraic metric).

The pose estimation methods presented by these works for noncentral catadioptric vision systems are non linear.

In this paper we study the pose estimation problem in the case of non-central 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 (proved in the text), 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.

Experimental tests performed with synthetic data and real images proved that the pose can be estimated using our linear algorithm accurately. In order to avoid numerical problems such as matrix ill-conditioning, quantization errors and bias, we used normalized coordinates both in the image and in 3D world data.

In the next section we present the details of our framework and some geometrical analytic tools are derived. The projection model in general catadioptric systems is then presented in section 3 and our linear method is implemented in section 4. Results of the experiments are presented in section 5 and section 6 concludes the work.

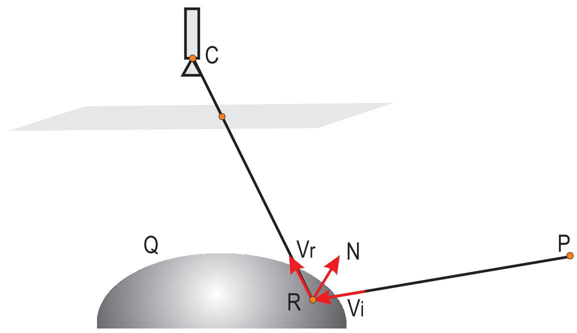


Figure 1. The light rays reflection and imaging in a catadioptric vision system.

2. Properties

In this section we present some notation conventions and mathematical results used throughout this paper.

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

Consider now 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 a quadric $\mathbf{Q}$ and 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 1 shows the reflection process and the notations adopted.

Some useful propositions are now presented.

Proposition 1 *Plane coordinates defined by three non collinear points can be expressed as a linear equation in the coordinates of one of the points.*

Proof:

Planes are defined by three points $\mathbf{U} = \begin{bmatrix} u_1 & u_2 & u_3 & u_4 \end{bmatrix}^T$, $\mathbf{V} = \begin{bmatrix} v_1 & v_2 & v_3 & v_4 \end{bmatrix}^\mathbf{T}$ and $\mathbf{W} = \begin{bmatrix} w_1 & w_2 & w_3 & w_4 \end{bmatrix}^\mathbf{T}$ (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 $\mathbf{M_\Pi} = \begin{bmatrix} \mathbf{X} & \mathbf{U} & \mathbf{V} & \mathbf{W} \end{bmatrix}$ with those three points and a generic point $\mathbf{X} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \end{bmatrix}^\mathbf{T}$.

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 $\mathbf{M_\Pi}$ must be zero. This gives us the expression of the plane in terms of the minors $\mathbf{D_{ijk}}$ of matrix $\mathbf{M_\Pi}$. It yields $\mathbf{\Pi} = \begin{bmatrix} \mathbf{D_{234}} & -\mathbf{D_{134}} & \mathbf{D_{124}} & -\mathbf{D_{123}} \end{bmatrix}^T$.

After rearranging the terms, the equation can the rewritten in the form $\mathbf{\Pi} = \mathbf{M}\mathbf{W}$, where the matrix $\mathbf{M}$ is symmetric. This equation is linear on $\mathbf{W}$. Matrix $\mathbf{M}$ is given by:

$$\mathbf{M} = \begin{bmatrix} 0 & u_3v_4 - u_4v_3 & -u_2v_4 + u_4v_2 & u_2v_3 - u_3v_2 \\ -u_3v_4 + u_4v_3 & 0 & u_1v_4 - u_4v_1 & -u_1v_3 + u_3v_1 \\ u_2v_4 - u_4v_2 & -u_1v_4 + u_4v_1 & 0 & u_1v_2 - u_2v_1 \\ -u_2v_3 + u_3v_2 & u_1v_3 - u_3v_1 & -u_1v_2 + u_2v_1 & 0 \end{bmatrix} \quad (1)$$

□

Lines are expressed by 4×4 Plücker matrices. For an Euclidean space, the absolute dual quadric is given by equation 2 where $\mathbf{I_3}$ represents the 3×3 identity matrix.

$$\mathbf{Q}^*_\infty = \begin{bmatrix} \mathbf{I_3} & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix} \quad (2)$$

Also useful is the angle between two planes. The cosine of the angle between planes $\mathbf{\Pi_A}$ and $\mathbf{\Pi_B}$ is given by equation:

$$cos\theta = \frac{\mathbf{\Pi_A}^T\mathbf{Q}^*_\infty\mathbf{\Pi_B}}{\sqrt{(\mathbf{\Pi_A}^T\mathbf{Q}^*_\infty\mathbf{\Pi_A})(\mathbf{\Pi_B}^T\mathbf{Q}^*_\infty\mathbf{\Pi_B})}} \quad (3)$$

In the next section we present the projection model of a general catadioptric system using a quadric mirror, deriving some constraints on the reflection point. These constraints are used in section 4 to estimate linearly the pose of the camera relative to the world reference frame.

3. Projection model in general catadioptric systems

In this section we present a projection model that can be applied to noncentral catadioptric systems made up by a quadric surface mirror and a perspective projection pinhole camera. The camera intrinsic parameters (focal length, principal point and skew parameter), the mirror and the pose of the camera relative to the mirror are assumed to be known.

Since the problem is how to project a 3D scene point into the image plane, the solution can be split into finding the reflection point $\mathbf{R}$ and projecting this point into the image plane. The second part of the solution is straightforward since the intrinsic parameters are known and thus the projection matrix as well.

Restrictions imposed on the reflection point

The solution of the projection problem is point $\mathbf{R}$. $\mathbf{R}$ 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 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 2).

   Proposition 2 *The reflection point $\mathbf{R}$ of a catadioptric system with quadric mirror $\mathbf{Q}$ is on the quadric $\mathbf{S}$, given by $\mathbf{S} = \mathbf{M}^T\mathbf{Q}^*_\infty\mathbf{Q}$, where $\mathbf{Q}^*_\infty$ is the absolute dual quadric and the 4×4 matrix $\mathbf{M}$ and the plane $\mathbf{\Pi_B}$ are defined by the 3D world point $\mathbf{P}$. The camera optical center $\mathbf{C}$ and the reflection point $\mathbf{R}$ are such that $\mathbf{\Pi_B} = \mathbf{MR}$.*

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

   The plane $\mathbf{\Pi_B}$ 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. The plane coordinates vector can be defined by a linear equation in the reflected point $\mathbf{R}$ as stated by $\mathbf{\Pi_B} = \mathbf{M}(\mathbf{P}, \mathbf{C}) \cdot \mathbf{R} = \mathbf{MR}$.

   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, $\mathbf{\Pi_A}$ and $\mathbf{\Pi_B}$, must be perpendicular. The angle between two planes is given by equation 3 and since we admit an Euclidean space, the absolute dual quadric for Euclidean transformations is given by expression 2.

   Since $\theta = \pi/2$ and substituting equations of the planes $\mathbf{\Pi_A}$ and $\mathbf{\Pi_B}$ into equation 3 it yields equation 4 which restricts the point $\mathbf{R}$ to belongs to a quadric given by $\mathbf{S} = \mathbf{M}^T\mathbf{Q}^*_\infty\mathbf{Q}$.

   $$\mathbf{\Pi_A}^T\mathbf{Q}^*_\infty\mathbf{\Pi_B} = 0 \Leftrightarrow \mathbf{R}^T\mathbf{Q}^T\mathbf{Q}^*_\infty\mathbf{MR} = 0 \Leftrightarrow \\ \Leftrightarrow \mathbf{R}^T\mathbf{M}^T\mathbf{Q}^*_\infty\mathbf{QR} = 0 \quad (4)$$

   Notice that matrix $\mathbf{S}$ is not symmetric as the generic quadric. 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.

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

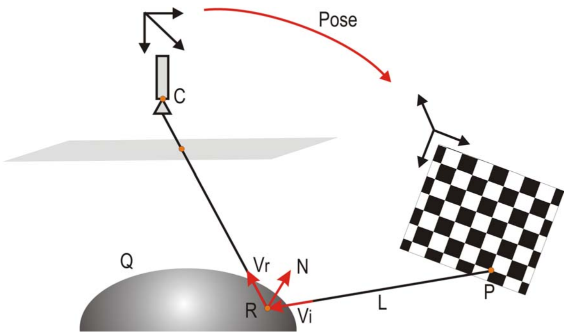


Figure 2. Pose transformation between the camera and world coordinate systems. The world coordinate system can be positioned in the scene object reference frame without loss of generality.

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.

Although it is possible to obtain a parametric representation of the intersection curve (a quartic) between quadrics $\mathbf{Q}$ and $\mathbf{S}$, depending on a single parameters, the solution is non linear and implicit, since the quartic must be searched for to find the point where incident and reflection angles are equal (restriction (3)). The intersection of quadrics could be computed by using either the Joshua Levin original method [15] or by the Dupont et al. alternative method [7].

4. Pose Estimation

In this section we develop a method to compute the pose of the catadioptric system in the world reference frame, assuming a calibrated catadioptric vision system, that is, 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. Similarly to other known methods to estimate pose [6, 16] the camera is assumed to be calibrated.

Using back projection the reflection point $\mathbf{R}$ is easily computed. The intrinsic parameters matrix $\mathbf{K}$ is inverted and the reflected ray emanated from the pinhole is intersected with the quadric to obtain the point $\mathbf{R}$.

Given a set of reflection points corresponding to some image pixels and its corresponding 3D world points whose local coordinates we know, the problem is then how to estimate the transformation matrix between the camera and local reference frames (made to coincide with the world reference frame without loss of generality) - see figure 2.

The basic idea is the expansion of the quadric matrix $\mathbf{S}$ elements given by proposition 2. Since $\mathbf{S} = \mathbf{M}^T\mathbf{Q}^*_\infty\mathbf{Q}$ and the reflection point $\mathbf{R}$ belongs to it, we can expand the equation $\mathbf{R}^T\mathbf{S}\mathbf{R} = 0$. The camera optical center $\mathbf{C}$ is known and the 3D world point $\mathbf{P}$ is given by $\mathbf{P} = \begin{bmatrix} p_1 & p_2 & p_3 & p_4 \end{bmatrix}^T$ in camera coordinates.

Expanding the equation and factorizing it in relation to the terms of the 3D point - p_i, it yields:

$$k_1p_1 + k_2p_2 + k_3p_3 + k_4p_4 = 0 \tag{5}$$

which is a 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. See appendix for explicit expressions.

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

$$\mathbf{P} = \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & t_{13} & t_{14} \\ t_{21} & t_{22} & t_{23} & t_{24} \\ t_{31} & t_{32} & t_{33} & t_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p_{w1} \\ p_{w2} \\ p_{w3} \\ p_{w4} \end{bmatrix} \tag{6}$$

where $\mathbf{P_W} = \begin{bmatrix} p_{w1} & p_{w2} & p_{w3} & p_{w4} \end{bmatrix}^T$ 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 5 the coordinates of the 3D world point given by equation 6, one obtains the following linear equation in the elements of the transformation matrix:

$$\begin{aligned} &k_1p_{w1}t_{11} + k_1p_{w2}t_{12} + k_1p_{w3}t_{13} + k_1p_{w4}t_{14} + \\ &+k_2p_{w1}t_{21} + k_2p_{w2}t_{22} + k_2p_{w3}t_{23} + k_2p_{w4}t_{24} + \\ &+k_3p_{w1}t_{31} + k_3p_{w2}t_{32} + k_3p_{w3}t_{33} + \\ &+k_3p_{w4}t_{34} = -k_4p_{w4} \end{aligned} \tag{7}$$

Each image point whose local 3D coordinates are known provides a different instantiation of equation 7. Using as many points as possible (a minimum of 12 points are needed) to enhance the robustness to noise, an overconstrained system is constructed and its least squares solution, if exists, is the pose of the camera in the world.

It is then possible to compute the least squares solution that best fits the observations. The solution given by the normal equations is $\mathbf{x} = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{b}$. A more robust estimator can be also 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 6 is affine instead of rigid, additional constraints are imposed to the problem to enforce orthogonality. There are several approaches to enforce $\mathbf{T_{rot}}^T\mathbf{T_{rot}} = \mathbf{I}$, where $\mathbf{T_{rot}}$ is the upper 3×3 submatrix of $\mathbf{T}$ 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 proscrustes solution in [18, 19].

As mentioned by Hartley [13], 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, the 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.

5. 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 Plucker coordinates so that their norm is unitary too.

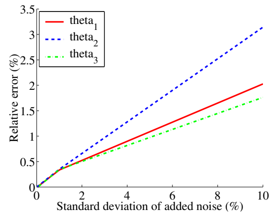
In tests with synthetic data, it is important to understand how the noise affects and degrades the pose estimate. We then performed Monte Carlo tests, repeating the estimation when random gaussian noise with zero mean was applied to each of the inputs separately, and choosing the median value for the 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 perspective.

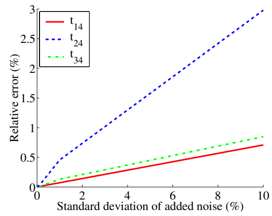
Figure 3 shows the results obtained for the hyperbolic mirror. The x-axis scale is linear. The standard deviation of the input gaussian noise added to variables varied from 10^{-7} (in percentage of the true value) to 10% of the true value (which is a considerable perturbation added to the inputs). Two input variables are tested: 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

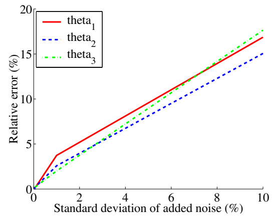
Figure 3. Error analysis tests with **simulated** data using a pinhole camera and a **hyperbolic mirror**. Noise was added separately to the input data of the algorithm: image points and 3D data points. Graphics (a) and (b) plot the error for respectively the euler angles and translation components when 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.



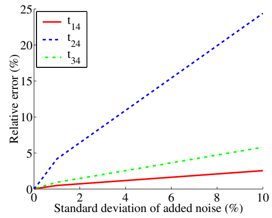
(a) θ_1, θ_2 and θ_3 with noise added to 3D points



(b) t_{14}, t_{24} and t_{34} with noise added to 3D points



(c) θ_1, θ_2 and θ_3 with noise added to image points



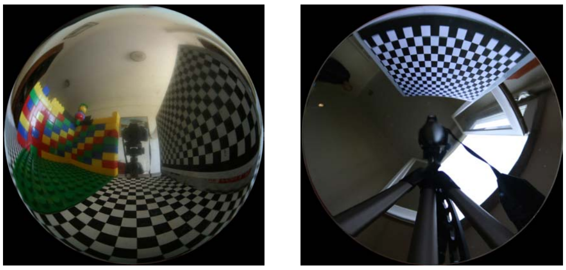
(d) t_{14}, t_{24} and t_{34} with noise added to image points

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 the positioning of the camera relative to the quadric mirror. Figure 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 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 [5]. 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 a method developed by the authors to calibrate catadioptric vision systems with quadric mirrors [11] was used. As a result the correspondence between pixels in the image and directions in space was estimated. Since the ground truth pose transformation was not



(e) Spherical mirror (f) Hyperbolic mirror

Figure 4. Real images used to estimate the pose of the camera in the world reference frame.

Table 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 our and Chen and Chang algorithms. The results are also compared before and after the nonlinear refinement performed to enhance accuracy.

	Before Ref.		After Ref.	
	Ours	**Chen&Chang**	**Ours**	**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
t_{14}	-101.138	-60.934	-57.905	-57.905
t_{24}	164.540	177.569	176.490	176.490
t_{34}	-336.094	-365.963	-366.805	-366.805

available we used an alternative method to compare our results with. Chen and Chang [6] algorithm was then applied to our input data. This method uses as input data the correspondences between a pixel and a direction 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 algorithm can then be compared.

We also evaluated the robustness of the method by computing the poses in two different positions of the catadioptric system relative 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 results obtained for the pose transformation (Euler rotation angles and translation elements) are presented in tables 1 and 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 the pose displacement and the ground truth motion

Table 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 our and Chen and Chang algorithms. The results are also compared before and after the nonlinear refinement made to enhance accuracy.

	Before Ref.		After Ref.	
	Ours	**Chen&Chang**	**Ours**	**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
t_{14}	-84.809	-82.974	-83.034	-83.034
t_{24}	126.133	126.138	126.895	126.895
t_{34}	179.138	171.421	173.559	173.559

Table 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.

	θ_1	θ_2	θ_3	t_14	t_24	t_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

are presented in table 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 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.

6. 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 relative 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 3). These constraints are linear in the coordinates of the 3D point projected in image.

The algorithm was compared to the Chen and Chang [6] algorithm to estimate the pose and the results are similar both with real and synthetic images. 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 contribution is the proof that linear algebraic methods can be applied to those extremely non-linear cameras, with non-central projection.

Appendix

Equation 5 is linear in the 3D point coordinates. Its coefficients k_i are given by the following equations:

$$\begin{aligned} k_1 = &q_{31}c_4r_1r_2 + q_{32}c_4r_2^2 + q_{33}c_4r_2r_3 + q_{34}c_4r_2r_4- \\ &- q_{21}c_4r_1r_3 - q_{22}c_4r_3r_2 - q_{23}c_4r_3^2 - q_{24}c_4r_3r_4+ \\ &+ q_{21}c_3r_1r_4 - q_{31}c_2r_1r_4 + q_{22}c_3r_2r_4 - q_{32}c_2r_2r_4+ \\ &+ q_{23}c_3r_3r_4 - q_{33}c_2r_3r_4 + q_{24}c_3r_4^2 - q_{34}c_2r_4^2 \end{aligned} \tag{8}$$

$$\begin{aligned} k_2 = &-q_{31}c_4r_1^2 - q_{32}c_4r_1r_2 - q_{33}c_4r_1r_3 - q_{34}c_4r_1r_4+ \\ &+ q_{11}c_4r_1r_3 + q_{12}c_4r_2r_3 + q_{13}c_4r_3^2 + q_{14}c_4r_3r_4- \\ &- q_{11}c_3r_1r_4 + q_{31}c_1r_1r_4 - q_{12}c_3r_2r_4 + q_{32}c_1r_2r_4- \\ &- q_{13}c_3r_3r_4 + q_{33}c_1r_3r_4 - q_{14}c_3r_4^2 + q_{34}c_1r_4^2 \end{aligned} \tag{9}$$

$$\begin{aligned} k_3 = &q_{21}c_4r_1^2 + q_{22}c_4r_1r_2q_{23}c_4r_1r_3 + q_{24}c_4r_1r_4- \\ &- q_{11}c_4r_1r_2 - q_{12}c_4r_2^2 - q_{13}c_4r_2r_3 - q_{14}c_4r_2r_4+ \\ &+ q_{11}c_2r_1r_4 - q_{21}c_1r_1r_4 + q_{12}c_2r_2r_4 - q_{22}c_1r_2r_4+ \\ &+ q_{13}c_2r_3r_4 - q_{23}c_1r_3r_4 + q_{14}c_2r_4^2 - q_{24}c_1r_4^2 \end{aligned} \tag{10}$$

$$\begin{aligned} k_4 = &-q_{21}c_3r_1^2 + q_{31}c_2r_1^2 - q_{22}c_3r_1r_2 + q_{32}c_2r_1r_2- \\ &- q_{23}c_3r_1r_3 + q_{33}c_2r_1r_3 - q_{24}c_3r_1r_4 + q_{34}c_2r_1r_4+ \\ &+ q_{11}c_3r_1r_2 - q_{31}c_1r_1r_2 + q_{12}c_3r_2^2 - q_{32}c_1r_2^2+ \\ &+ q_{13}c_3r_2r_3 - q_{33}c_1r_2r_3 + q_{14}c_3r_2r_4 - q_{34}c_1r_2r_4- \\ &- q_{11}c_2r_1r_3 + q_{21}c_1r_1r_3 - q_{12}c_2r_2r_3 + q_{22}c_1r_2r_3- \\ &- q_{13}c_2r_3^2 + q_{23}c_1r_3^2 - q_{14}c_2r_3r_4 + q_{24}c_1r_3r_4 \end{aligned} \tag{11}$$

where the camera optical center is $\mathbf{C} = \begin{bmatrix} c_1 & c_2 & c_3 & c_4 \end{bmatrix}$.

References

[1] S. Baker and S. Nayar. A theory of catadioptric image formation. In *ICCV*, pages 35–42, Bombay, 1998.

[2] J. Barreto and H. Araújo. Geometric properties of central catadioptric line images. In *ECCV*, 2002.

[3] J. Barreto and H. Araújo. A general framework for selecting world coordinate systems in perspective and catadioptric imaging applications. *IJCV*, April 2004.

[4] M. Born and E. Wolf. *Principles of Optics*. Pergamon Press, 1965.

[5] J.-Y. Bouguet. Camera calibration toolbox for matlab. Internet web page: http://www.vision.caltech.edu/bouguetj/calib_doc/.

[6] C.-S. Chen and W.-Y. Chang. On pose recovery for generalized visual sensors. *PAMI*, 26(7), 2004.

[7] L. Dupont, S. Lazard, S. Petitjean, and D. Lazard. *Towards the Robust Intersection of Implicit Quadrics*, volume 704, chapter 5, pages 59–68. Kluwer Academic Publishers, 2002.

[8] M. Fischler and R. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *ACM Graphics and Image Processing*, 24(6):381–395, 1981.

[9] X.-S. Gao, X.-R. Hou, J. Tang, and H.-F. Cheng. Complete solution classification for the perspective-three-point problem. *PAMI*, 25(8):930–943, 2003.

[10] C. Geyer and K. Daniilidis. Catadioptric projective geometry. *IJCV*, 45(3):223–243, 2001.

[11] N. Gonçalves and H. Araújo. Low-cost method for the estimation of the shape of quadric mirrors and calibration of catadioptric cameras. *Optical Engineering*, 46(7):1–12, July 2007.

[12] M. Grossberg and S. Nayar. A general imaging model and a method for finding its parameters. In *ICCV*, Vancouver, Canada, July 2001.

[13] R. Hartley. Minimizing algebraic error in geometric estimation problems. In *ICCV*, 1998.

[14] R. Horaud, B. Conio, O. Leboulleux, and B. Lacolle. An analytic solution for the perspective 4-point problem. *CVGIP*, 47(1):33–44, 1989.

[15] J. Levin. A parametric algorithm for drawing pictures of solid objects composed of quadric surfaces. *Communications of the ACM*, 19(10):555–563, 1976.

[16] D. Nister. A minimal solution to the generalised 3-point pose problem. In *CVPR*, 2004.

[17] R. Pless. Using many cameras as one. In *CVPR*, 2003.

[18] P. Schönemann. A generalized solution of the orthogonal proscrustes problem. *Psychometrika*, 31(1), March 1966.

[19] P. Schönemann. Fitting one matrix to another under choice of a central dilation and a rigid motion. *Psychometrika*, 35(2), June 1970.

[20] H. Stewénius, D. Nistér, M. Oskarsson, and K. Åström. Solutions to minimal generalized relative pose problems. In *OMNIVIS'05*, Beijing, China, 2005.

[21] P. Sturm and S. Ramalingam. A generic camera calibration concept. In *ECCV04*, pages 1–13, Prague, May 2004.

[22] Y. Wu and Z. Hu. Pnp problem revisited. *Journal of Mathematical Imaging and Vision*, 24:131–141, 2006.

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