Abstract

This paper addresses the problem of rigid motion estimation and 3D reconstruction in vision systems where it is possible to recover the incident light ray direction from the image points. Such systems include pinhole cameras and catadioptric cameras. Given two images of the same scene acquired from two different positions, the transformation is estimated by means of an iterative process. The estimation process aims at having corresponding incident rays intersecting at the same 3D point. Geometrical relationships are derived to support the estimation method. Furthermore, this paper also addresses the problem of the mapping from 3D points to image points, for non-central catadioptric cameras with mirror surfaces given by quadrics. The projection model presented can be expressed in a non-linear equation of only one variable, being more stable and easier to solve than the classical Snell’s law. Experiments with real images are presented, by using simulated annealing as estimation method.

1. Introduction

3D reconstruction from images has been extensively studied in the past and several methods exist providing results with good accuracy (e.g. [10, 27]). In the case of catadioptric images, 3D reconstruction is more complex. In this paper we study the problem of reconstruction in the case of non-central catadioptric systems.

In the case of central catadioptric systems the epipolar geometry has already been derived and studied. Geyer and Daniilidis [4, 6] have studied these issues as well as Svorba and Pajdla [24]. In this case, it is possible to use most of the results obtained for 3D reconstruction with pinhole cameras – epipolar geometry, bundle adjustment for example. Several algorithms have been proposed and implemented [5, 25].

For non-central catadioptric vision systems there is a viewpoint surface instead of a single viewpoint which implies other solutions for the correspondence and reconstruction rather than the epipolar geometry. The study of the viewpoints surfaces called caustics has been performed by Nayar et al. [26] and the geometrical properties of generalized catadioptric systems were studied [9, 17, 18]. In [8] the mirror surface is recovered for a catadioptric system with quadric mirrors. In the field of structure from motion several works have been done [7, 20], using two or more cameras or using one moving camera. New camera designs have also been proposed to optimize the extraction of specific kinds of information from images [14, 22]. The combination of several cameras or a camera with several mirrors has also been studied in order to build a non central generalized vision system [7, 19].

The projection model for non-central cameras (that use mirrors) mapping 3D world points into the image plane can be derived using only the specular reflection laws - Snell’s law - [1, 12] - so that the image point is a function of the 3D world point, the mirror surface and the location and orientation of the optical axis of the camera. However, when projecting a 3D world point, the main effort is finding out the reflection point in the mirror surface. Furthermore, the projection is highly non linear in several variables resulting in a difficult problem to solve. This issue is important to image formation, computer graphics applications, and simulation purposes.

In this paper we address the problem of 3D reconstruction as well as the estimation of rigid motion. This problem has been addressed, for example, by using bundle adjustment. However, this paper makes no central projection assumption and thus there is no closed loop or implicit projection function that could be used in the bundle model.
Therefore, our aim is to study the 3D reconstruction and motion estimation for a catadioptric vision system, assuming a generalized camera as in [9]. We show that this framework can be extended to any catadioptric vision system. As an example of the general nature of the method we also address the case of a pinhole camera.

We also address the problem of 3D to image mapping in non-central catadioptric cameras using quadric mirror surfaces. The method presented derives restrictions on the coordinates of the reflection point, and the solution is obtained by solving a nonlinear equation in only one parameter.

2. Problem statement

The reflection through a generic curved mirror (see figure 1) is in accordance to the well known Snell’s law ([1, 12]). The reflection law is given by equation 1 where \( \mathbf{V}_i \) is the incident light ray, \( \mathbf{V}_r \) is the reflected light ray and \( \mathbf{N} \) is the normal vector to the mirror surface.

\[
\mathbf{V}_i = \mathbf{V}_r - \mathbf{N} \times \mathbf{N} \quad (1)
\]

The two problems that have to be addressed are (1) how to project a 3D world point in the image plane and (2) given an image point how can we recover the incident direction (direction of \( \mathbf{V}_i \)) - back projection (notice that the reflected ray \( \mathbf{V}_r \) can be computed from the image as long as the intrinsic parameters are known). The problem of how to recover the actual 3D point that was projected is even harder.

The solution for the first problem stated - the projection model - is actually equivalent to solving the problem of estimating the reflection point \( \mathbf{R} \). However, even if the analytical expression of the specular surface is known, it is difficult to estimate the coordinates of the reflection point \( \mathbf{R} \) due to the non-linearity of the estimation of the surface normal.

To back project a light ray let us consider an image point. The back projection solution is the incident light ray direction \( \mathbf{V}_i \). Since we assume that the coordinates of the camera optical center are known as well as the image point (the intrinsic parameters are assumed to be known), the reflected ray \( \mathbf{V}_r \) can be computed easily.

This problem is easier to solve if the mirror surface is known, since in that case intersecting the surface with the reflected ray is straightforward. Next the normal vector at the reflection point is estimated and equation 1 can be used to estimate \( \mathbf{V}_r \).

However, if the mirror surface is not known, the solution can only be approximated. In [8] an algorithm to solve nonlinear equations has been used to estimate not only the reflection point \( \mathbf{R} \) but also the coefficients of the mirror analytical expression. The algorithm considered the mirror to be a ruled quadric and it assumed the knowledge of the coordinates of 3D points. The reference frame was located in the center of the quadric. As we shall see in the next section, when a general quadric is used and there is no assumption on the origin of the coordinate system, the estimates of the parameters are very difficult to obtain with good accuracy since the nonlinear equation to solve has several unknowns.

The 3D reconstruction of a scene point from the incident direction can only be performed using additional information.

In the next section we propose an analytical method to project scene to image points in a catadioptric system composed by a perspective camera and a quadric surface mirror. We assume the knowledge of the quadric and of the 3D coordinates of the scene point.

In section 4 we propose a method to recover the 3D coordinates of a scene point from a pair of images as well as the camera rigid motion.

All points are represented by their 4 × 1 vector of homogeneous coordinates \( X = [x_1  \ x_2  \ x_3  \ x_4] \) such that its corresponding cartesian point is \( x = [x_1/x_4  \ x_2/x_4  \ x_3/x_4] \). Quadrics are represented by a symmetric 4 × 4 matrix \( Q \) and any point \( X \) is on the quadric surface if and only if \( X^T Q X = 0 \).

3. Projection model

In this section we present a projection model that can be applied to non-central catadioptric cameras composed by a quadric surface mirror and a perspective projection pinhole camera. The camera intrinsic parameters, the quadric and the pose of the camera relative to the mirror are assumed to
be known.

3.1. Some geometrical properties

Planes are defined by three points $A$, $B$ and $C$ (generating points).

Consider a plane $\Pi$ and define an auxiliary matrix $W = [X \ A \ B \ C]$ with those three points and a generic point $X$. It can be easily shown that the vector of the coordinates of the plane can be expressed as a linear combination of one of its generating points given by equation 2, since the determinant of matrix $W$ must be zero.

\[ \Pi = MC \tag{2} \]

where

\[
M = \begin{bmatrix}
-a_3 a_4 + a_3 b_4 & a_3 b_4 - b_3 b_4 & a_3 b_4 - b_3 b_4 & a_3 b_4 - b_3 b_4 \\
-a_3 a_4 + a_3 b_4 & a_3 b_4 - b_3 b_4 & a_3 b_4 - b_3 b_4 & a_3 b_4 - b_3 b_4 \\
-a_3 a_4 + a_3 b_4 & a_3 b_4 - b_3 b_4 & a_3 b_4 - b_3 b_4 & a_3 b_4 - b_3 b_4 \\
-a_3 a_4 + a_3 b_4 & a_3 b_4 - b_3 b_4 & a_3 b_4 - b_3 b_4 & a_3 b_4 - b_3 b_4
\end{bmatrix}
\]

Another relevant property is relative to the angle between two planes. Although the sign of the angle cannot be defined in a projective space $P^3$, as pointed out by Stolfi [23], the cosine of the angle is well defined and is given by equation 4, where $Q_\infty$ is the absolute dual quadric.

\[ \cos \theta = \frac{\Pi_T^T Q_\infty^* \Pi_T}{\sqrt{(\Pi_T^T Q_\infty^* \Pi_A)(\Pi_T^T Q_\infty^* \Pi_B)}} \tag{4} \]

3.2. Restrictions imposed by specular reflection

The solution of the problem is point $R$. $R$ is the reflection point on the mirror surface that projects the 3D point $P$ into the image plane. For such point the following restrictions must be imposed:

1. $R^T QR = 0 \rightarrow$ the point is on the quadric of the mirror surface.
2. $R^T SR = 0 \rightarrow$ the point is on the quadric given by $S = M^T Q_\infty^* Q$ (proposition 1).

**Proposition 1** The reflection point $R$ of a catadioptric camera with quadric mirror $Q$ is on the quadric $S$, given by $S = M^T Q_\infty^* Q$, where $Q_\infty^*$ is the absolute dual quadric, the $4 \times 4$ matrix $M$ and the plane $\Pi_B$ are defined by the 3D world point $P$, the camera optical center $C$ and the reflection point $R$ are such that $\Pi_B = MR$.

**Proof:** Let us consider two concurrent planes: $\Pi_A$ and $\Pi_B$. $\Pi_A$ is the tangent plane to the quadric $Q$ in the reflection point $R$. Its representation is given by $\Pi_A = QR$. The plane $\Pi_B$ is the plane defined by three points: the camera optical center $C$, the 3D point $P$ and the reflection point $R$ in the mirror surface. Using equation 2 the plane coordinates vector can be defined by a linear equation in the reflected point $R$ as stated by $\Pi_B = M(P, C) \cdot R = MR$.

Since the normal to the quadric is perpendicular to the tangent plane and must be on the plane defined by the three points $C$, $P$ and $R$, then the two planes, $\Pi_A$ and $\Pi_B$, must be perpendicular. The angle between two planes is given by equation 4.

Since $\theta = \pi/2$ and substituting equations of planes $\Pi_A$ and $\Pi_B$ into equation 4 it yields equation 5 which tells that the point $R$ belongs to a quadric given by $S = M^T Q_\infty^* Q$.

\[ \Pi_A^T Q_\infty^* \Pi_B = 0 \Leftrightarrow R^T Q^T Q_\infty^* M R = 0 \Leftrightarrow \]
\[ \Leftrightarrow R^T M^T Q_\infty^* Q R = 0 \tag{5} \]

Notice that matrix $S$ is not symmetric as the generic quadric. However, without loss of generality, matrix $S$ can be substituted by another matrix whose entries are related by $S_{ij} = -0.5 S_{ij} + 0.5 S_{ji}$. With this change the quadric remains the same and its representing matrix becomes symmetric. \( \square \)

3. The incidence and reflected angles are equal.

Since the angles of incidence and reflection are equal, the angle between the incident and reflected rays and the tangent line to the quadric in the reflected point are also equal. After some simplifications, one obtains expression 6 for the third restrictions (see [23] for details in the angle between two lines). In this expression, the line $l_{BQ}$ is the tangent line to the quadric that pass in $R$ and that is on the plane defined by $C$, $P$ and $R$.

\[ \frac{\text{dir}(l_{BC})^T \text{dir}(l_{BQ})}{\sqrt{\text{dir}(l_{BC})^T \text{dir}(l_{BC})}} = \frac{\text{dir}(l_{RP})^T \text{dir}(l_{BQ})}{\sqrt{\text{dir}(l_{RP})^T \text{dir}(l_{RP})}} \tag{6} \]

The directions $l_{BC}$ and $l_{RP}$ are defined by the join of two points [23]. The line $l_{BQ}$ is defined as the intersection of the planes $\Pi_A$ and $\Pi_B$ so its direction $\text{dir}(l_{BQ})$ is computed from the Plücker matrix of equation 7.

\[ L_{BQ} = \Pi_A \Pi_B^T - \Pi_B \Pi_A^T = \]
\[ = Q R \cdot (M R)^T - M R \cdot (Q R)^T = \]
\[ = Q R R^T M^T - M R R^T Q^T \tag{7} \]
3.3. Computing the reflection point \( R \)

Given the three restrictions imposed to the reflection point \( R \), the problem is now how to find that point. The first and second restrictions are much similar since they restrict the point \( R \) to be on quadric \( Q \) (restriction (1)) and to be also on quadric \( S \) (restriction (2)). This is obviously the problem of finding the intersection of those two quadrics (a curve in space). Since the third restriction constrains the point so that the incident and reflection angles are equal, point \( R \) must be located in the intersection curve.

The general method for computing an explicit parametric representation of the intersection between two quadrics is due to Joshua Levin [15,16]. However, the parametric representation of this method is hard to compute and is less reliable due to the high number of irrational numbers needed. Hence some alternative methods can be used in the computation of the intersection curve [2,3].

The parametric curve given by the intersection algorithm is a function of only one parameter, say \( \lambda \). Let us represent the parameterized curve by the \( 4 \times 1 \) vector \( X(\lambda) \). Although nonlinear, the curve can be searched for the point where incident and reflected angles are equal. Let us call \( \lambda_0 \) to the value of the parameter that solves equation 6. The resulting reflection point is given by \( R = X(\lambda_0) \).

This method finds the reflection point \( R \) in a non-central catadioptric vision system presents a major advantage over the method of using explicitly the Euclidean expressions of the mirror (quadric mirrors - conic section) and its normal vector expressions. This advantage is the fact that once intersected the quadrics \( Q \) and \( S \), the solution is given by a nonlinear equation in only one parameter. Instead of that, if the explicit Euclidean expressions were used, the problem was to solve a nonlinear equation in several unknowns without good guesses for the initial search point. This is very important for the accuracy of the solution.

4. 3D Reconstruction and Rigid Motion Recovery

4.1. General Framework

To develop our framework for this section several relationships involving lines, planes and points in \( P^3 \) were derived.

The basic idea presented here is the following. Let us consider an arbitrary black box camera such that one is able to calculate the incident direction of the light rays that are projected into a given image point (see figure 2). The incident direction is represented by the Plücker matrix \( L \). There is no assumption of central projection.

If the camera moves to another position (this motion is represented by rigid transformation \( T \)) and tracking an image point along both frames, it is possible to calculate the incident ray in both positions: \( \ell_1 \) and \( \ell_2 \). Each of the Plücker matrices of the lines \( \ell_1 \) and \( \ell_2 \), \( L_1 \) and \( L_2 \), are in the proper reference frame, before and after the motion. Since both are related by the rigid transformation \( T \), the Plücker line matrix \( L_2 \) is given by \( T^T L_2 T \) in the initial one. The 3D point reconstructed is the intersection of both lines \( L_1 \) and \( T^T L_2 T \). However, the transformation \( T \) is not known.

Given the lines of the incident directions \( L_1 \) and \( T^T L_2 T \) in the initial reference frame, one is able to intersect them and recover the point \( P \). The problem then becomes how to estimate a transformation \( T \) such that both lines intersect in point \( P \).

Our proposal is to split the problem in two parts. We propose first the estimation of a transformation \( T \) that meets a particular condition and then the intersection of the lines in 3-space to estimate point \( P \).

4.2. Geometric relationships

In this subsection some geometric relationships in projective 3-space \( P^3 \) are derived (see [11,21] for background support).

**Proposition 2** Given the lines \( \ell_1 \) and \( \ell_2 \) in \( P^3 \), represented by their corresponding Plücker matrices \( L_1 \) and \( L_2 \), they intersect each other if and only if \( L_1 L_2^T L_1 = 0 \), where \( L_2^* \) is the dual representation of the \( L_2 \) Plücker matrix.

**Proof:** To prove that the condition is necessary we assume that the lines \( \ell_1 \) and \( \ell_2 \) intersect. Let us consider an arbitrary plane \( \Pi_\alpha \). If the plane contains \( \ell_1 \) one has \( L_1 \Pi_\alpha = 0 \) and then nothing can be concluded about the matrix \( L_1 L_2^T L_1 \). If the plane contains the intersection point but not line \( \ell_1 \) nothing either can be concluded since one has \( L_1 \Pi_\alpha = X_{1\alpha} \) which is the point of intersection of the plane \( \Pi_\alpha \) and \( \ell_1 \) and since it is also the intersection of \( \ell_1 \) and \( \ell_2 \) and then is on \( L_2 \), one has \( L_2^* X_{1\alpha} = L_2^* L_2 \Pi_\alpha = 0 \).

However, if the plane \( \Pi_\alpha \) does not contain any of the two lines nor their intersection point, \( X_{1\alpha} = L_1 \Pi_\alpha \) is the intersection point of plane \( \Pi_\alpha \) with line \( \ell_1 \). This point isn’t on the line \( \ell_2 \) and then \( \Pi_\alpha = L_2^* X_{1\alpha} = L_2^* L_1 \Pi_\alpha \) is the plane defined by line \( \ell_2 \) and point \( X_{1\alpha} \). And by hypothesis,
the two lines intersect each other and then the line $\ell_1$ is on the plane $\Pi_b$ (notice that if $\ell_1$ isn’t on the plane $\Pi_b$ the only common point with this plane is $X_{1a}$ which is not on the line $\ell_2$ by definition). We thus have $L_1\Pi_b = L_1L_2L_1\Pi_a = 0$.

Since the plane $\Pi_a$ is arbitrary one thus conclude the thesis, that is $L_1L_2^2L_1 = 0$.

The counterpart should be now proved. Assume that $L_1L_2^2L_1 = 0$ and then let us try to prove that the lines intersect. Consider again an arbitrary plane $\Pi_c$ not containing any of the two lines nor their intersection point. Multiplying the plane $\Pi_c$ in the right of both sides of the condition we obtain $L_1L_2^2L_1\Pi_c = 0$, where $L_1\Pi_c$ represents the intersection point of the line $\ell_1$ and the plane $\Pi_c$. Let’s say $X_c$. Since this point is not on $\ell_2$, $L_2^2X_c = L_2^2L_1\Pi_a$ represents the plane defined by $\ell_2$ and $X_c$. say $\Pi_d$. $L_1\Pi_d = 0$ and then line $\ell_1$ is on this plane. Since both $\ell_1$ and $\ell_2$ are on the same plane $\Pi_d$, they intersect each other. This proves the sufficiency of the condition. □

**Corollary 1** Consider two intersecting lines $\ell_1$ and $\ell_2$. If line $\ell_1$ is transformed by a general point transformation, given by $H$, the transformed line $\ell_1'$ and $\ell_2$ no longer intersect in the general case.

**Proof:** By proposition 2 one has $L_1L_2^2L_1 = 0$. Line $\ell_1$ is transformed and the new line is given by $L_1' = HL_1H^T$. Computing the intersection condition of proposition 2, it yields $C = L_1'L_2^2L_1' = HLL_1H^T L_2HLL_1H^T$. In the general case $C \neq 0$ and so the two lines no longer intersect. □

**Proposition 3** The intersection point of two arbitrary 3-space intersecting lines $\ell_1$ and $\ell_2$ in $\mathbb{P}^3$ is given by $P = L_1L_2A$, where $A$ is an arbitrary point not belonging to the plane defined by $\ell_1$ and $\ell_2$.

**Proof:** Consider plane $\Pi_{2A}$ defined by an arbitrary point $A$ and $\ell_2$ so that $\Pi_{2A} = L_2^2A$. Since $A$ does not belong to the plane defined by $\ell_1$ and $\ell_2$, the intersection of $\ell_1$ with $\Pi_{2A}$ is point $P$, given by $P = L_1\Pi_{2A} = L_1L_2^2A$. □

4.3. Rigid Motion Estimation and 3D Reconstruction

To recover the rigid transformation matrix $T$ we propose the use of the geometric relations of the previous subsection. After recovering the incident direction from a single point in the image in both reference frames, the problem becomes the recovery of the transformation $T$ between both coordinate systems.

Any nonlinear minimization algorithm can be applied to this problem in order to estimate the transformation $T$ subject to the intersection condition (proposition 2). However, there are multiple solutions guaranteeing the intersection between both lines, as long as the transformed line is on the pencil of planes defined by the line in the first frame. The goal is to find out which rigid transformation $T$ should be chosen so that both lines intersect in the correct point $P$.

The estimation of such transformation is not possible if only one or two pairs of lines are used. However, if we use three or more arbitrary points the only transformation that assures the intersection of all the pairs of lines is the rigid transformation $T$ we are seeking. Figure 3 represents a catadioptric vision system in two positions and the corresponding incident directions intersecting in three 3-space points.

The second and last step in the overall algorithm is the estimation of the 3-space points that project on the image. Since all pairs of the estimated incident directions intersect, the 3D points to reconstruct are the intersections of the lines. The expression for the intersection is given by proposition 3.

We presented a method to estimate the rigid motion of a camera after a general rigid transformation from an initial position. This method also estimates the 3D coordinates of the points. The main contribution of this method is the introduction of a way to recover both the pose and 3D points when no projection model exists (in which case bundle adjustment or any other known method could be used). In the next two sections we parameterize the method in two particular configurations: the pinhole camera and the catadioptric vision system with quadric shape mirrors.

4.4. Pinhole Camera

The application of this framework to a pinhole camera can serve as a reference to compare the implementation with other camera models.

The camera projection model used is $p_{image} = K \cdot Proj \cdot P_{3D}$, where $K$ is the intrinsic parameters matrix. Given $K$ and any image point tracked along at least two frames,
one has to recover the incident direction of light rays for all frames. The line that represents the incident direction is recovered in the local reference frame, not taking into account the motion between frames.

To invert the projection model we arbitrate the value of one of the coordinates, say $Z = 1$, and one can thus obtain a point in the incident direction. Expressing the other two coordinates in relation to the third, it yields the following expression for the point $D$ in the incident light ray:

$$
\begin{align*}
X &= \frac{y - z_0 - k z^2 z_0}{f} \\
Y &= \frac{v - z_0}{f} \\
Z &= 1
\end{align*}
\Rightarrow D \sim \begin{bmatrix} y - z_0 - k z^2 z_0 \\ v - z_0 \\ 1 \\ 1 \end{bmatrix} \quad (8)
$$

The coordinates of two points are then known from the incident direction, the origin of the reference frame ($O = [0 \ 0 \ 0 \ 1]$) and point $D$. The corresponding Plücker matrix that we search for, representing the incident direction, is given by:

$$
L_i = O \cdot D^T - D \cdot O^T = \begin{bmatrix} 0 & 0 & 0 & -d_1 \\ 0 & 0 & 0 & -d_2 \\ 0 & 0 & 0 & -d_3 \\ d_1 & d_2 & d_3 & 0 \end{bmatrix} \quad (9)
$$

where $d_i$ is the $i$-th coordinate of point $D$.

This recovers the incident direction given the intrinsic parameters and one image point tracked along the frames.

### 4.5. Catadioptric Camera

The advantages of a wide field of view justify the use of catadioptric vision systems in many applications. Although some particular configurations of the camera in relation to the mirror produces a central projection system, in the general case the central projection constraint is relaxed to enable other important characteristics (e.g. zooming). The relaxation of the central projection constraint has many implications in the model and in the algorithms used. There is no general projection model since each point has its own viewpoint (see [26]). This turns the method presented in this paper into an important tool to reconstruct the scene.

The catadioptric vision system used is one of the most common - a curved mirror given by a quadric and a pinhole camera. Consider a quadric mirror represented by the $4 \times 4$ matrix $Q$ so that $X^T Q X = 0$ for all points $X$ on its surface. To calculate the incoming light ray that reflects to the camera, one should first calculate the reflected ray direction and intersect it with the mirror surface.

Using the same reasoning applied to the pinhole camera, the direction $L_i$ must be intersected with the quadric.

**Figure 4. Geometric construction used to find out the incident direction.**

**Proposition 4** The intersection points $R$ of the full rank quadric $Q$ with the line $\ell$ given by the Plücker matrix $L$ is given by the eigenvectors of matrix $LQ$.

**Proof:** The tangent plane $\Pi_N$ to the quadric surface in a point $R$ is $\Pi_N = QR$. Since the intersection of line $\ell$ with the quadric is the same point of intersection of line $\ell$ with the plane $\Pi_N$, it yields $R = L \Pi_N = LQR$ which is the same to say that $R$ is an eigenvector of $LQ$. Since matrix $L$ has rank 2 and matrix $Q$ has full rank, matrix $LQ$ has two eigenvalues that correspond to the two points of intersection. $\square$

**Proposition 5** The normal line to the quadric given by its $4 \times 4$ matrix $Q$ at point $R$ is given by $L_N = RR^T Q^T Q^*_\infty - Q^*_\infty QRR^T$, where $Q^*_\infty$ is the dual absolute quadric.

**Proof:** The tangent plane to the quadric through $R$ is given by $\Pi_N = QR$ and the direction of this plane is given by $\text{dir}(\Pi_N) = Q^*_\infty \Pi_N = Q^*_\infty QR$. Since the direction of a plane also represents the intersection of its normal line with the plane at infinity [23], the normal line is the joint of the points $R$ and $\text{dir}(\Pi_N)$, given by $L_N = R d \text{dir}(\Pi_N)^T - \text{dir}(\Pi_N) R^T$ or $L_N = RR^T Q^T Q^*_\infty - Q^*_\infty QRR^T$.$\square$

To find out the incident direction of the incoming ray the reflection law is used. It is known that the angle between incident and reflected rays with the normal vector is equal. The geometric construction represented in figure 4 is used to calculate one point in the incoming ray. As shown in figure 4, $G$ is the point on the incoming ray such that the angle between line $\ell_i$ ($L_i$, join of points $R$ and $G$) and the normal $\ell_N$ ($L_N$, join of points $R$ and $E$) is the same as the angle between the reflected ray $\ell_r$ ($L_r$, join of point $R$ and $C$) and the normal line $\ell_N$. $E$ is any point of the normal line yielding $G = \mu E - C$. The value of the parameter $\mu$ is such that the angles of incidence and reflection are equal. It can be solved with a second degree linear equation.

Having calculated the value of $\mu$, the incident direction is computed with points $R$ and $G$. The method presented in section 4.1 can then be used.
5. Experiments

Images of the real world with structured indoor scenes are used to test the rigid motion estimation and 3D reconstruction models presented. The images were taken by a catadioptric system composed by a 640 × 480 resolution pinhole camera with zoom and an hyperbolic mirror. Before applying the algorithms proposed, the system was calibrated in two steps: pinhole camera calibration and mirror calibration. The pinhole camera parameters were calibrated using Heikkilä and Silven’s algorithm [13]. The mirror parameters and the distance from pinhole to mirror were calibrated using a chess board panel, and by iterating the parameters until the corners form the real chess board. Figure 5 shows one shot taken by the system described.

A similar catadioptric system was modelled and some synthetical data were used to test both rigid motion and 3D reconstruction. The projection in images were determined by using the Snell’s Law and the projection model presented in section 3.

To test the method presented in section 4 using synthetical data we first recovered the rigid motion transformation matrix and then the 3D scene points. The nonlinear estimation algorithm used was simulated annealing, since it uses a stochastic jump to iterate to the actual configuration. Although the minimization algorithm is not compared to others, we experienced good results with simulated annealing. Two different motions are used: only translation along the coordinate axes and rotation and translation along the optical axis. Table 1 presents some of the results obtained compared with the ground truth values.

The results of the experiments with real images are shown in table 2. The motion transformation is recovered for each pair using the corners of a chess board panel. Although 3D reconstruction is performed, since the correct 3D coordinates of the points are not known, the estimated values are not shown. Five different motion pairs were used, including only translation along the reference axis, only rotation and both translation and rotation.

We can observe from the experiments with both synthetical or real images, that the method nearly converges to the actual transformation and that the errors in 3D coordinates of the points are low as expected.

6. Conclusion and directions

This paper presents a projection model for non central catadioptric image formation. Since no central projection assumption is made, the problem of mapping 3D world points to image plane is difficult to solve and we show that it is possible to model this problem as a non linear equation in only one variable. The paper also presents a model to estimate the rigid motion transformation of a moving camera from frame to frame and subsequently make 3D reconstruction.

The results with real images show that it is possible with the theory presented to compute accurately the motion of the camera from frame to frame and then to estimate the 3D positions of corresponding points in both images. The experiments however use a minimization method that doesn’t run in real time due to the stochastic nature of the optimization.

The enhancement of the method performance and the extension of the theory to non calibrated cameras and mirrors are the next problems to be addressed.

References


<table>
<thead>
<tr>
<th>Ground truth</th>
<th>Estimated</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\theta_x$</td>
<td>0.0</td>
</tr>
<tr>
<td>$\theta_y$</td>
<td>-0.003</td>
</tr>
<tr>
<td>$\theta_z$</td>
<td>1.571</td>
</tr>
<tr>
<td>$t_x$</td>
<td>0.0</td>
</tr>
<tr>
<td>$t_y$</td>
<td>0.0</td>
</tr>
<tr>
<td>$t_z$</td>
<td>1.571</td>
</tr>
</tbody>
</table>

Table 1. Experiments with synthetical data. Rigid motion estimation ($\theta_i$ are rotation angles in radians and $t_i$ are translations in mm) and 3D reconstruction of space coordinates $X$, $Y$ and $Z$.


