Abstract

In this paper we address the problem of sphere localization from a single image taken with a noncentral catadioptric camera. We propose a method for determining both the radius and the position of an unknown sphere from a single, catadioptric image. The method can find its application in the field of robotic vision, especially in mobile robots playing soccer in RoboCup contests, in order to improve robot capabilities related to playing with a flying ball.

Recently, a method for sphere reconstruction from single image taken with a noncentral, axial-symmetric, catadioptric camera has been proposed. In an axial symmetric catadioptric cameras, the pinhole of the camera is placed on the mirror axis. Though axial symmetric cameras help to simplify the geometrical treatment of the problem, they are difficult to set up since they require a precise alignment, usually hard to check. In this paper we deal with the general case of off-axis catadioptric cameras, with the camera pinhole placed in a general position with respect to the mirror. We devise a simple geometrical method by which we determine both the position of a sphere and its radius from its apparent image contour. Since our approach is based on coplanar viewing rays, it has a wider applicability with respect to the previous method, as it relaxes the constraint on camera position, i.e. it does not require a precise alignment, and the constraint on mirror axial symmetry, i.e. it can be applied to a wider class of mirrors. Some preliminary experiments both on simulated and real image are also presented.

1. Introduction

This paper is related to sphere localization; in particular we study the determination of both the radius and the 3D position of a sphere from a single image. Using a standard perspective camera, if the radius is known, a sphere can be localized from its apparent contour image. Consider a sphere of radius \( R \) and let \( D \) be the (unknown) distance between the sphere center and the camera viewpoint. Consider now the two points of the contour image of the sphere for which the angle between the associated viewing rays is maximum: if \( \alpha \) is the semi-angle between these two rays, then it yields that \( \sin \alpha = R/D \), which can be solved to determine \( D \). Therefore the sphere can be easily localized.

This task becomes challenging when no a priori information about the sphere is given. In this work we propose a method to tackle this problem by using a non-central catadioptric camera. We show that under broad conditions we can determine both the radius and position of a sphere from a single catadioptric image.

Detection, localization and tracking of spherical objects have been widely applied to video sport analysis for automatic extraction of information [7], analysis of matches and tactics [9], video indexing and summarization [11], and verification of controversial referee decisions [18]. In sports like soccer, tennis, volleyball and basketball, a ball is an important feature to extract from images.

Many efforts have been done in estimating the 3D position of the ball from a video sequence. Matsumoto et al. [14] used multiple fixed cameras and, once the ball is detected by template matching, the position is estimated by epipolar line constraints between multiple cameras. Reid and Zisserman [18] proposed a method for accurate metrology from uncalibrated sequences. The method uses two images acquired simultaneously from different viewpoints and determines the vertical projection of the ball onto the ground plane by using the homography induced by the ground plane. Hence the ball position is estimated from two images while the size of the ball is supposed to be known.

Other approaches exploit the trajectory of the ball to determine its position [10]. However, all these methods determine the position of the ball using many images taken from sequences of images or from different viewpoints. The size of the ball (i.e. the radius) is supposed to be known (each sport has a fixed and standard size for the ball).

Sphere localization find application in the field of mobile robotics, where the localization of a sphere, i.e. the ball, is a crucial task for robots playing soccer in RoboCup contests [22]. In order to detect and localize a ball, many approaches have been employed. Scaramuzza et al. in [19]
used a stereo vision system to give an optimal estimation of the 3D position of the ball and the other objects in the field of play.

The approach with a single frontal perspective camera is also widely exploited: the most common techniques for ball detection rely on color information and are based on fast color segmentation of the image to detect and track objects [4].

Catadioptric cameras are also used since they enlarge the field of view of the camera and allow the robot to see all the objects around. Ye et al. [24] used a catadioptric camera to estimate the angular position of the ball, assuming it is on the ground. Sun et al. [20] performed the Cylinder Projection of the catadioptric image by warping the image onto a cylinder; then they tracked the ball in order to get the angular direction, assuming the ball to be on the ground.

Catadioptric stereo or multi-catadioptric images have been also investigated: Maeda et al. [12] proposed a vision system composed of three catadioptric cameras while Ying and Hu [25] developed a motion estimate method based on two successive catadioptric images of three or more spheres.

However, for all these methods the 3D position of the ball can not be computed without knowing the radius of the ball (the size of the ball is standardized in each RoboCup League), or making other assumptions, such that all relevant objects are located on the ground of the field [20], or by using team coordination methods in order to integrate in a SLAM framework the sensing information coming from other teammates [23].

Recently a method for determining both the radius and the position of a unknown sphere has been proposed [1]. The method exploits simple geometrical properties of non-central, axial symmetric catadioptric cameras in order to retrieve the pose and the radius of the sphere from a single image. In this paper we propose a new method that can be applied to a general class of noncentral catadioptric cameras by relaxing (i) the constraint of alignment of the camera viewpoint with the mirror axis and (ii) the constraint of axial symmetry for the mirror.

The paper is structured as follows: in Section 2 some preliminary notions on catadioptric cameras and their geometry are introduced. In Section 3 the addressed problem is formulated and the assumptions that underlie the method are presented. In Section 4 we describe the geometric aspects of our method and an algorithm for the determination of radius and distance of a sphere is presented. In Section 5 some preliminary experimental results that validate the method are presented.

2. Preliminary notions

A catadioptric camera is generally constituted by a curved mirror placed in front of a perspective camera. In a catadioptric camera, the viewing ray coming from a scene point is specularly reflected by the mirror surface according to the reflection laws, before it goes through the camera viewpoint and crosses the image plane. Catadioptric cameras are attractive because of the possibility to employ them in omni-directional vision.

A catadioptric camera is calibrated, if the viewing ray associated to each image point is known.

If all viewing rays concur at a same point (called “center”), the camera is said to be central. Baker and Nayar [2] derived the complete class of central catadioptric cameras which preserve the single viewpoint constraint. Central cameras can be obtained by placing a hyperbolic mirror, or an elliptical mirror, in front of a perspective camera, such that the viewpoint of the camera is on one of the foci of the mirror surface. In this way, the center, where all viewing rays concur, is on the other focus. Hence any calibrated central camera reduces to a perspective camera.

Noncentral cameras, cameras whose viewing rays are not all concurrent [21], can be realized, e.g., by linear push-broom cameras [8], cross-slit cameras [6], catadioptric cameras [15].

With respect to central cameras, noncentral cameras remove some degeneracies: e.g., the viewing surface, i.e., the union of the viewing rays, of a straight line in the 3D space is not planar. This fact has can be exploited to localize a straight line in the 3D space from a single image [5].

On the other hand, some difficulties arise in the image-based characterization of curved (self-occluding) surfaces. In particular, using a central camera, the plane tangent to a curved surface at a point on the contour generator can easily be derived from the image: it is the backprojection of the tangent to the apparent contour. Using a noncentral camera, since the viewing rays can be skew, the image-based characterization of the plane tangent to a curved surface is more difficult.

In the sequel, we will focus on single-image catadioptric cameras, i.e., cameras such that through any space point external to the mirror surface there is only one viewing ray. It can be shown that if the mirror surface is convex the camera is single-image.

An axial-symmetric catadioptric camera consists of an axial-symmetric mirror and a perspective camera, whose viewpoint $O$ is on the symmetry axis of the mirror. The mirror surface of an axial-symmetric catadioptric camera can be obtained by rotating a planar curve, called profile, about the symmetry axis. On the other hand, in a off-axis catadioptric camera, the camera viewpoint is placed in a general position w.r.t. the mirror surface. Though axial symmetric cameras are attractive because of their simple geometry, they are difficult to set up since they require accurate and hard to check alignments. In this paper we consider general catadioptric cameras.
3. Problem formulation

Now we formulate the addressed problem. A noncentral catadioptric camera placed in front of a convex mirror in a general position is given. A sphere of unknown radius is placed at an unknown position in the 3D space. The sphere is visible from the catadioptric camera, but it can be partially occluded or not entirely visible from the camera. The apparent image contour of the sphere is extracted. The addressed problem is to determine both the radius and the position of the sphere.

In a recent work [1] a similar problem was addressed and a method for determining both the radius and the position of an unknown sphere was proposed. The method, however, relies only on axial symmetric catadioptric cameras. In this work we relax this constraint and we proposed a more general method based on less restrictive assumption. In the following we assume that:

- the camera is calibrated and the mapping between any pixel and its associated viewing ray is known; moreover the camera is not oblique [16], i.e. the viewing rays are not all skew;
- the mirror surface is supposed to be convex, so that the catadioptric system is single-image. While in [1] the mirror surface has to be convex and axial symmetric, we relax the latter constraint since the method can be applied even to non axial symmetric mirrors;
- the constraint on camera position w.r.t. the mirror surface is relaxed and the camera can be placed in a generic position, which allows a more flexibility in the design of the catadioptric sensor;
- The sphere has been detected in the image and the relevant contour points have been extracted.

In the following we analyze the geometry of general cameras and describe our general method.

4. General cameras

Relaxing the constraint on axial symmetry, the geometrical properties on which the previous reconstruction method is based no more yield. We propose a new method based on the idea of coplanar viewing rays. Since the camera is calibrated, checking if two viewing rays are coplanar is a relatively simple task. By exploiting \( N \geq 3 \) pairs of coplanar viewing rays we devised a simple and effective reconstruction method that holds for a wider class of noncentral catadioptric cameras.

Consider two viewing rays tangent to the sphere. If the viewing rays are coplanar, the plane \( \pi_c \) containing both rays intersects the sphere. The intersection is a circle, that is not generally a great circle. Consider now the plane \( \pi_b \) symmetric w.r.t. the two viewing rays (note that it contains the bisectrix of the two viewing rays and it is perpendicular to the plane \( \pi_c \)). From classical geometry, the plane \( \pi_b \) goes through the center of the sphere. Therefore, in order to univocally determine the center of the sphere at least \( 3 \) symmetry planes are required, i.e. at least \( 3 \) (distinct) pairs of coplanar viewing rays are required to determine the position of the center. The estimate of radius can be calculated straightforwardly as the distance of the viewing rays from the center.

The method we propose uses \( N \geq 3 \) pairs of coplanar viewing rays in order to robustly estimate the position of the sphere and its radius. The main steps of our algorithm are:

1. Starting from the contour points of the sphere, choose \( N \geq 3 \) pairs of coplanar viewing rays. In order to select a pair of coplanar viewing rays, for each contour point we calculate the associated viewing ray by backprojection. Then we consider \( N \) contour points and for each of them we find the corresponding contour point, whose associated viewing rays is coplanar.

2. For each pair, determine the symmetry plane which also contain the bisectrix of the two rays.

3. Intersect all the symmetry planes and find the center of the sphere. Intersecting \( N \geq 3 \) planes requires to solve an overdetermined linear system of \( N \) equations in \( 3 \) unknown.

4. Once the center is determined, the radius of the sphere can be estimated as the mean distance of each considered viewing rays from the calculated center of the sphere.

While determining the viewing ray associated to a contour point is straightforward from the system calibration, finding two coplanar viewing rays requires to match the viewing rays associated to the contour points. Given a viewing ray \( r \) associated to a point \( p \) on the image contour \( c \) of the sphere, we devised the following procedure in order to find a corresponding coplanar viewing ray:

a) for each point \( p_i \in c \) calculate the relevant (signed) minimum distance between the ray \( r_i \) (associated to \( p_i \)) and \( r \). If the distance is equal to zero, then the rays intersect in one point, hence they are coplanar. Otherwise they are skew.

b) The viewing rays associated to points of \( c \) close to \( p \) will possibly have a distance value near zero, but they are not coplanar. In order to choose more robustly a reliable coplanar viewing ray, we consider the trend of the distance value along the contour \( c \) and then we consider only the viewing rays whose distance is around the zero-crossing point of the trend.
c) Interpolate the contour points associated to the found viewing rays with a cubic spline and then find the contour point \( p_x \) whose associated viewing ray \( r_x \) has minimum distance from \( r \).

d) The viewing rays \( r_x \) associated to \( p_x \) is the viewing ray coplanar to \( r \).

The proposed method exploits simple geometrical properties and relationships that can be applied to noncentral catadioptric cameras placed in a general position w.r.t. to mirror.

5. Experimental Results

In order to validate the proposed method we performed some preliminary experiments. We used a catadioptric camera based on a conical mirror which is relatively easy to manufacture.

In an off-axis camera based on conical mirror, meridian lines could be exploited in order to speed up the search for coplanar viewing rays. A meridian line is a line through the cone apex contained in the cone surface. The viewing rays associated to points on the mirror surface lying on the same meridian are coplanar. Hence (the image of) a meridian line can be exploited to find pairs of contour points whose associated viewing rays are coplanar. However, this technique relies on an accurate localization of the cone apex, which may be noisy and introduces unwanted error in the estimation process. In our experimental activities we do not use this method but we follow the method detailed in Section 4.

In the following we present some experimental results obtained both with simulated and real images.

5.1. Simulated Data

In order to validate and evaluate the accurateness of the proposed method we perform some experiments using synthetic images rendered with POV-Ray [17], an open source ray tracing engine. Since the shape and the position of the camera w.r.t. the mirror is known from the 3D model as well as the camera parameters, the method was applied to estimate both radius and distance.

The most critical aspect of the proposed method is the robustness of the solution calculated by plane intersection. Initial experimental activities showed that the \( N \) planes are usually almost parallel and, instead of intersecting in a single point, they tend to share a common supporting line, on which the actual center lie. Indeed, we find out that the overdetermined system of \( N \) equation used to estimate the center has no maximum rank (i.e. one of the eigenvalue is close to 0), thus affecting the reliability of the solution. In order to overcome this issue and better fit the position of the center, we then apply a further optimization step by minimizing the sum of two terms. The first term considers the residuals of the mean radius calculated w.r.t. all the viewing rays, as the center of the sphere should lie at the same distance from each ray. The second term we considered constrains the center to lie on all the \( N \) planes. Hence the minimization has the form:

\[
\arg \min_{\mathbf{c} \in \mathbb{R}^3} \left[ w \sum_{i=1}^{N} (d - d(r_i, \mathbf{c})) + (1 - w) \sum_{i=1}^{N} (\pi_i \mathbf{c} + q_i) \right]
\]

where \( d(r_i, \mathbf{c}) \) is the euclidean distance of the center \( \mathbf{c} \) from the \( i \)-th viewing ray \( r_i \), \( d \) is the mean distance of \( \mathbf{c} \) from all the viewing rays (i.e. the mean radius), \( \pi_i \mathbf{c} + q_i \) is the algebraic distance of \( \mathbf{c} \) from the \( i \)-th plane \( (\pi_i, q_i) \), with \( \pi_i \in \mathbb{R}^3 s.t. \| \pi_i \| = 1 \). The sum of the two terms can be weighted with \( w \in (0, 1) \), so one can better fit the position w.r.t. radius and viceversa.

As for all 3D reconstruction methods, we were interested in verifying the reconstruction accuracy of the method w.r.t. the distance of the sphere. Moreover we also compared the performance of the method according to the camera position. We model the mirror as a highly reflective cone. Then we used two spheres of radius 4.5cm and 11cm respectively as test cases. We placed them at an initial position and then we moved them along the viewing ray passing through the center, so that the projected image size decreases. Figure 2 and Figure 3 show the reconstruction errors plotted against the distance obtained with an axial camera (Figure 2.a and Figure 3.a, respectively) and a slightly misaligned camera (Figure 2.b and Figure 3.b respectively). For each distance value, we run 10 trial choosing randomly 10 pairs of coplanar viewing rays and then we report the median value of estimated radii and distances. The errors grow larger and larger as the sphere goes far away from the mirror because the small baseline gets very small compared to the distance. However the errors are reasonable within distances up 10-15 times the radius of the sphere. Moreover in real applications when the sphere is tracked along different frame one could apply the proposed method to initially estimate the radius and position of the sphere; then the estimated value of the radius can be used in the optimization process in the

Figure 1. The symmetry planes used to estimate the center of the sphere. Most of the planes are almost parallel and share a common supporting line on which the center lie. A further optimization step is needed to improve localization.
subsequent frames to get a more robust and accurate estimate position of the frame.

5.2. Real images

We also performed some preliminary experiments with real images in order to validate the method with a catadioptric camera based on a conical mirror. We remark that the method has a more general validity and can be applied to a broad class of convex mirror.

For these preliminary experimentations we followed a rough two-step calibration, following the method proposed by Mashita et al. [13] because of its simplicity. The method exploits the mirror boundaries and estimates the mirror posture w.r.t. the camera, using the ellipse in image corresponding to the boundary of the mirror. First, the intrinsic camera parameters is estimated using the Matlab Calibration Toolbox [3]. Then the external circumference of the conical mirror is localized w.r.t. the camera. The accuracy of this localization technique relies on sufficient accuracy of the mirror shape and it may introduce some errors.

In order to validate the method, we compared the performances of the method in [1] with our method. We manually aligned the camera viewpoint to the mirror axis and then we applied the two methods. For the sake of these experiments we manually extract the sphere contour since the detection

<table>
<thead>
<tr>
<th>Image</th>
<th>Ground truth</th>
<th>Method [1]</th>
<th>Proposed method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tennis</td>
<td>3.2 20</td>
<td>4.28% 6.9%</td>
<td>1.89% 2.64%</td>
</tr>
<tr>
<td>Soccer</td>
<td>10.9 49</td>
<td>2.41% 5.8%</td>
<td>1.03% 0.88%</td>
</tr>
<tr>
<td>Tennis 1</td>
<td>3.2 25</td>
<td>n.a. n.a.</td>
<td>7.55% 8.95%</td>
</tr>
<tr>
<td>Tennis 2</td>
<td>3.2 27</td>
<td>n.a. n.a.</td>
<td>6.64% 9%</td>
</tr>
</tbody>
</table>

Table 1. Localization results for real image. $R$ and $D$ are the ground truth measures (in cm) of radius and distance respectively; $e_R$ and $e_D$ are the percentage estimation errors.

Figure 2. Experiments with a sphere of radius $4.5\text{cm}$: the graphs show the reconstruction errors (%) of radius $R$ and distance $D$ at different distances.

(b) Camera slightly misaligned.

Figure 3. Experiments with a sphere of radius $11\text{cm}$: the graphs show the reconstruction errors (%) of radius $R$ and distance $D$ at different distances.

of the ball in image is beyond the scope of this paper; however color information could be exploited in order to automatically detect the sphere in the image and then extract the contours. Figure 4 reports the images used to compare the performances, while Table 1 reports the relevant estimation errors. The estimation errors are lower than [1] in both images. We also performed some experiments with the camera placed in a generic position (Figure 5). The reconstruction accuracy of our method (method [1] is not applicable in these cases) is reported in the lower part of Table 1. Our method performed well since it relies on a general model of noncentral camera, while the other method relies on the axial symmetry of the system which is simpler to treat but more difficult to achieve.
The method to a more general imaging model. The method on various kind of mirror surfaces, and at extending presented. Mental results that validate the proposed method are also apparent image contour is described. Preliminary experiments are aimed at experimenting the proposed method on a broad class of noncentral catadioptric cameras since it relaxes the constraint of axial symmetry, which is physically hard to obtain and difficult to check. The geometric properties that underlies the reconstruction method are presented and a simple but effective algorithm for retrieving the radius and the position of the sphere from its apparent image contour is described. Preliminary experimental results that validate the proposed method are also presented.

Ongoing works are aimed at experimenting the proposed method on various kind of mirror surfaces, and at extending the method to a more general imaging model.

6. Conclusions

In this paper we propose a method for determining both the radius and the position of an unknown sphere in space from a single catadioptric camera. The method can be applied to a broad class of noncentral catadioptric cameras since it relaxes the constraint of axial symmetry, which is physically hard to obtain and difficult to check. The geometric properties that underlies the reconstruction method are presented and a simple but effective algorithm for retrieving the radius and the position of the sphere from its apparent image contour is described. Preliminary experimental results that validate the proposed method are also presented.

Ongoing works are aimed at experimenting the proposed method on various kind of mirror surfaces, and at extending the method to a more general imaging model.

References


Figure 4. A catadioptric image of a tennis ball (a) and of a soccer ball (b) with the relevant extracted contour points

(a) Tennis 1 (b) Tennis 2

Figure 5. A catadioptric image of a tennis ball with the camera in an off-axis position w.r.t. the mirror