Catadioptric camera calibration based on RANSAC

Fuqing Duan\textsuperscript{1,2}, Mingquan Zhou\textsuperscript{1}, Fuchao Wu\textsuperscript{2}, and Xiaoming Deng\textsuperscript{3,2}

\textsuperscript{1} College of Information Science and Technology, Beijing Normal University, Beijing 100875, P.R China
\textsuperscript{2} National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100080, PR China
\textsuperscript{3} Virtual Reality Laboratory, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080, P.R. China

\textbf{Abstract.} In this paper, we propose a simple method for calibrating the effective focal length of central catadioptric cameras. This method uses the constraint that the projections of any three collinear space points on the viewing sphere should be on a great circle. We assume there are some inliers among the detected image points. If only three image points of a space line are well chosen, the constraint is effective. RANSAC is used for robust estimation. This approach needs no conic fitting, which is hard to accomplish and highly affects the accuracy of the calibration. Experiments on simulated and real data show the proposed method is robust and effective.

\textbf{Keywords:} Camera calibration, catadioptric camera, RANSAC

\section{Introduction}

In many computer vision applications such as robot navigation, surveillance, teleconferencing and virtual reality, it would be convenient if the imaging system could have a large field of view. One effective way to enhance the field of view is to combine mirrors with conventional cameras, which is called a catadioptric imaging system \cite{1}. In catadioptric systems, a single effective viewpoint is highly desirable due to its superior and useful geometric properties \cite{2,5,6}. A catadioptric system with a unique viewpoint is called a central catadioptric system. The complete class of central catadioptric systems was presented by Baker and Nayar \cite{1}. They introduced that a central catadioptric system can be built by setting a parabolic mirror in front of an orthographic camera, or a hyperbolic, elliptical, planar mirror in front of a perspective camera, where the single viewpoint constraint can be fulfilled via a careful alignment of the mirror and the camera. Now the calibration of central catadioptric cameras has been an active research field \cite{7-11}.

Geyer and Daniilidis \cite{6} proposed a unified sphere model for describing central catadioptric cameras, under which some algorithms \cite{3,7,10,11,13} were proposed for calibrating central catadioptric cameras. Ying and Hu \cite{13} applied the projections of lines or spheres to calibrate central catadioptric cameras. Wu and Duan
presented a linear algorithm based on the projections of space points. Barreto and Araujo [2] studied the projective invariant properties of catadioptric images of space lines and showed that any central catadioptric camera can be fully calibrated from an image of three or more lines. Ying and Zha [14] presented some identical projective geometric properties of central catadioptric images of lines and spheres, and applied these properties to calibration. Nearly all these approaches need conic fitting since a line in space is projected to a conic in a central catadioptric image, and the accuracy of the calibration highly depends on the accuracy of the conic fitting. The conic is called the line image. In general, only a small segment of the conic is visible in the catadioptric image due to the partial occlusion, which makes the conic estimation hard to accomplish. Wu et al [12] presented a calibration method of no conic estimation, but it was mainly for para-catadioptric cameras. In this paper, we propose a nonlinear algorithm for partially calibrating any central catadioptric cameras, which needs no conic estimation. Since some parameters of the camera can be obtained by prior information about scene or system configuration, we only estimate the effective focal length. The algorithm is based on the straight line projection constraint and Random Sample Consensus (RANSAC)[4].

Section 2 reviews the unified sphere model given by Geyer and Daniilidis. Section 3 describes the proposed method. Experimental results are reported in Section 4, and followed are some conclusions in Section 5.

2 Central catadioptric camera

Geyer and Daniilidis [6] showed that the central catadioptric imaging process was equivalent to the following two-step mapping by a sphere (see Fig. 1):

1. Under the viewing sphere coordinate system $O-xyz$, a 3D point $X = [x, y, z]^T$ is projected to a point $X^s$ on the unit sphere centered at the viewpoint $O$ by $X^s = [x/r, y/r, z/r]^T, r = ||X||$.

2. The point $X^s$ on the viewing sphere is projected to a point $m$ on the image plane $\Pi$ by a pinhole camera through the perspective center $O^c$. The image plane is perpendicular to the line going through the viewpoints $O$ and $O^c$, and it is also called the catadioptric image plane.

In this camera system, the optical axes of the pinhole camera is the line $O^cO$, and thus its principal point is the intersection, $p = [u_0, v_0, 1]^T$, of the line $O^cO$ with the image plane $\Pi$. The distance from point $O$ to $O^c$, $\xi = ||O - O^c||$ is called the mirror parameter, which determines the mirror used in the central catadioptric camera. The mirror is a paraboloid if $\xi = 1$, an ellipsoid or a hyperboloid if $0 < \xi < 1$, and a plane if $\xi = 0$. The details can be found in [6]. In this paper, we assume $0 < \xi \leq 1$, i.e. do not consider the case of plane mirror. Let the intrinsic matrix of the pinhole camera be

$$K = \begin{bmatrix} rf & s & u_0 \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (1)$$
Where $f$ is the effective focal length; $r$ is the aspect ratio; $p = [u_0, v_0, 1]^T$ is the principal point; $s$ is the parameter describing the skew of the two image axes. Then, the catadioptric image of a space point $X$ is

$$m = \lambda K[I, \xi e] \begin{bmatrix} X^s \\ 1 \end{bmatrix} = \lambda K(X^s + \xi e), \quad (2)$$

with $\lambda$ being a scalar, $I$ the $3 \times 3$ identical matrix, and $e = [0, 0, 1]^T$.

In the catadioptric camera calibration, there are totally six parameters \{r, f, s, u_0, v_0, \xi\} to be determined.

### 3 Calibrating Algorithm

In [10], we have derived the projections of space points as follows:

**Proposition:** Let $m$ be the catadioptric image of a space point $X$. Then, under the pinhole coordinate system the projection of the point $X$ on the viewing sphere can be expressed as

$$X^s = \frac{\xi(1 + \sqrt{1 + \tau \eta})}{\eta} K^{-1} m. \quad (3)$$

In the case of paraboloid mirror, we have

$$X^s = \frac{2}{\eta} K^{-1} m. \quad (4)$$

where $\tau = (1 - \xi^2)/\xi^2$, $\eta = m^T \varpi m$, $\varpi = K^{-T}K^{-1}$.

According to the unified sphere model, the spherical projections of any three collinear space points, \{X^s_j, j = 1, 2, 3\}, should be on a great circle. That is, the
viewing sphere center O is on the plane defined by the three points. Thus, we have
\[
f(\tau, \varpi) = \det[X_s^1 - O, X_s^2 - O, X_s^3 - O].
\] (5)

Although central catadioptric cameras have six parameters to be determined, sometimes we can get several of them by some prior information about scene or system configuration, For example, the principal point can be set with the image center, or determined from four collinear points \cite{11} or the bounding ellipse of the catadioptric image \cite{9}, the skew can be set as zero, and the aspect ratio as one. Here assume that the principal point, skew factor, the aspect ratio and mirror parameter are all known, and we only estimate the effective focal length \( f \). Then one line image is enough.

Previous approaches using lines have to estimate the conic curves corresponding to the line images. However, it is very difficult to get the conic curves correctly even for a lower noise level due to partial occlusions. Thus constraints derived from the conic curves are largely biased. Instead we use the nonlinear constraint \( (5) \) from a line image. We assume there are some inliers among the detected image points, that is, they are without noise or with small amounts of noise. If only the three image points are well chosen, the constraint is effective and the focal length can be estimated. Therefore, the problem of the parameter estimation is transferred into chosen of three image points of a space line and further determining a great circle of the viewing sphere.

RANSAC is an iterative method for robust fitting of a model from a set of observed data which contains many outliers. We use RANSAC to determine the great circle from the detected points on a line image, consequently estimate the focal length. In each iteration step of RANSAC, a great circle is determined by optimizing the focal length of the camera.

**Calibration Algorithm Based on RANSAC**

Given the catadioptric image of a space line, and denote the detected points on the line image as \( D \). Assume all other camera parameters are known, estimate the effective focal length \( f \):

Step 1. Randomly select three image points from \( D \);

Step 2. By the constraint \( (5) \), construct a cost function \( F(\tau, \varpi) = f(\tau, \varpi)^2 \).

Step 3. Given an initial value of the effective focal length \( f \), minimize the cost function using some nonlinear technique, e.g. Levenberg-Marquart algorithm.

Step 4. Compute the projections \( P \) on the viewing sphere of the point set \( D \) using \( (1) \), and determine a great circle from the three image points.

Step 5. Compute the sum of the distances from the projections \( P \) to the plane containing the great circle, and return to step1.

Step 6. Perform K random samplings, and choose the parameter corresponding to the least distance sum as the final output.

**Remark 1:** In step 5, assume the unit normal vector of the plane is \( n \), then the distance from a spherical point \( p \) to the plane is \( d = |p^T n| \).

**Remark 2:** Assume the ratio of the inliers is \( \rho \) and the probability that there is a good sampling among the K samplings is \( z \), then \( [4] \)
Initial Estimation

In [13], some intrinsic parameters of the camera can be determined from the bounding ellipse of the catadioptric image. In fact, since both the eccentricity of the mirror and the field of view (FOV) are usually known, we can determine all intrinsic parameters from the bounding ellipse.

Assume the bounding ellipse is denoted as

\[ ax^2 + 2bxy + cy^2 + 2dx + 2xy + f = 0. \]  

(7)

The authors of [6] give the following estimation:

\[
\begin{align*}
    r &= \sqrt{-\frac{b^2}{a^2} + \frac{c}{a}}, \\
    s &= -\frac{b}{a}, \\
    u_0 &= \frac{bc-ce}{bd-ae}, \\
    v_0 &= \frac{bd-ad}{bd-ae},
\end{align*}
\]

(8)

Now we describe how to estimate \( f \) and \( s \). It’s well known that the bounding ellipse is the projection of the mirror boundary, which is a circle, and the optical axis is perpendicular to the plane containing the circle and goes through the center of the circle.

Let the distance from the optical center \( O^c \) to the center \( O^h \) of the circle be \( h \) and the radius of the circle be \( q \). Then, the projection of the point \( X_0 = (q, 0, h)^T \) is \( m_0 = (\frac{r_h}{h^\top} + u_0, v_0, 1)^T \), which is one intersection of the bounding ellipse and the line \( v = v_0 \) on the image plane. As Fig. 2 shows, \( q \) and \( h \) can be decided from FOV, i.e. \( q = \sin \theta, h = \xi + \cos \theta, \theta = \frac{\text{FOV}}{2} \). Then, from the point \( m_0 \) and the equation (8), we can determine all intrinsic parameters.

4 Experimental results

We use simulated and real data to evaluate the performance of our algorithm.

4.1 Sensitivity analysis

The intrinsic parameters of the simulated catadioptric cameras are \((r, f, s, u_0, v_0) = (500, 500, 0, 512, 384)\). The mirror parameter \( \xi \) is 0.9 or 1. One catadioptric line image is randomly generated by choosing one unit normal vector corresponding to a great circle on the viewing sphere. To simulate actual conditions, we choose 50 points on a one-third portion of the entire circle, and project these points to the catadioptric image plane. Gaussian noise with zero mean and standard deviation \( \sigma \) is added to each image point. The noise level \( \sigma \) is varied from 0 to 5 pixels with a step of one pixel. For each noise level, we perform 50 trails. In each trail, we keep all other intrinsic parameters unchanged except the focal
length, which is initialized by uniformly sampling in the interval [400, 600]. The means and standard deviations of the estimated focal lengths with respect to different noise levels are shown in Tab.1 and Tab.2 for $\xi = 0.9$ and $\xi = 0.1$ respectively. From the two tables, we can see that the estimation accuracy is degraded gracefully with respect to the noise level. The result is accurate with $\sigma = 0$, and is very close to the truth even with large noises. It shows that the proposed method is effective and robust.

Since it is possible that other intrinsic parameters are not accurate in real applications, here we consider the algorithm’s sensitivity to other parameters. The simulated data is the same as before. When adding Gaussian noise with the noise level $\sigma = \sigma_i$ to the image points, we also add noise with $2\sigma$ to the principal point and noise with $\sigma = 1$ to the skew. Usually the mirror parameter offered by manufacturers is nearly precise, so we add Gaussian noise with $0.01\sigma$ to the mirror parameter in each trail. The estimated results are shown in Tab.3 and Tab.4, corresponding to Tab.1 and Tab.2. Comparing these tables, we can see that for each mirror parameter, the means in the two tables are very close, while standard deviations in the latter are larger than the ones in the former. However, we still can conclude that the approach is robust to variations of the other intrinsic parameters.

**Table 1.** Means and standard deviations of the estimated focal lengths when adding Gaussian noise to Image points for the mirror parameter is 0.9

<table>
<thead>
<tr>
<th>$\sigma$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>500</td>
<td>500.12</td>
<td>500.56</td>
<td>499.98</td>
<td>501.17</td>
<td>502.83</td>
</tr>
<tr>
<td>standard deviation</td>
<td>0</td>
<td>1.65</td>
<td>3.14</td>
<td>4.56</td>
<td>6.25</td>
<td>9.36</td>
</tr>
</tbody>
</table>
Table 2. Means and standard deviations of the estimated focal lengths when adding Gaussian noise to Image points for the mirror parameter is 1

<table>
<thead>
<tr>
<th>σ</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>500</td>
<td>500.48</td>
<td>500.77</td>
<td>501.67</td>
<td>502.4</td>
<td>502.95</td>
</tr>
<tr>
<td>standard deviation</td>
<td>0</td>
<td>1.57</td>
<td>2.87</td>
<td>4.13</td>
<td>6.21</td>
<td>6.67</td>
</tr>
</tbody>
</table>

Table 3. Means and standard deviations of the estimated focal lengths when adding Gaussian noise to Image points and all other parameters for the mirror parameter is 0.9

<table>
<thead>
<tr>
<th>σ</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>500</td>
<td>500.66</td>
<td>500.3</td>
<td>500.31</td>
<td>501.57</td>
<td>503.51</td>
</tr>
<tr>
<td>standard deviation</td>
<td>0</td>
<td>5.41</td>
<td>7.01</td>
<td>8.42</td>
<td>10.4</td>
<td>13.37</td>
</tr>
</tbody>
</table>

Table 4. Means and standard deviations of the estimated focal lengths when adding Gaussian noise to Image points and all other parameters for the mirror parameter is 1

<table>
<thead>
<tr>
<th>σ</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean</td>
<td>500</td>
<td>500.06</td>
<td>501.45</td>
<td>503.62</td>
<td>502.54</td>
<td>502.93</td>
</tr>
<tr>
<td>standard deviation</td>
<td>0</td>
<td>5.73</td>
<td>7.0</td>
<td>10.4</td>
<td>9.6</td>
<td>14.15</td>
</tr>
</tbody>
</table>

4.2 Real data

The used catadioptric system consists of a perspective camera with a hyperbolic mirror. The mirror is designed by the Center for Machine Perception, Czech Technical University, its FOV is 217.2 degree, and the eccentricity of the hyperbolic mirror is 1.302, corresponding to $\xi = 0.966$. One image of an indoor scene is taken with a resolution of $2048 \times 1536$. Five lines are used and shown in Figure 3a. The image points of the lines and the mirror boundary are manually selected by using the software downloaded from [8]. The intrinsic parameters of the camera estimated from the bounding ellipse of the catadioptric image are $(r, f, s, u_0, v_0) = (501.62, 501.95, -0.5, 1017, 753)$. So the initial value of the focal length is 501.95. Since each line has different partial occlusion and noise
level, the distance sum in the step5 of the algorithm is computed from all image points of the five lines in order to improve the estimation accuracy. The estimated results using the five lines are 556.19, 559.33, 568.9, 551.43, 546.92 respectively, where the distance sum corresponding to the result from the first line is the minimal. The rectified image using the result is shown in Figure.3b. We can see the rectified lines are straight.

5 Conclusion

Nearly all approaches using lines need conic fitting, which is hard to accomplish and highly affects the accuracy of the calibration. In this paper, we propose a nonlinear algorithm for calibrating the effective focal length of central catadioptric cameras. This method is based on the straight line projection constraint and RANSAC, and needs no conic estimation. Experiments on simulated and real data show the proposed method is robust and effective.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (Grant No.60872127), China Postdoctoral Science Foundation funded project, and the K. C. Wong Education Foundation, Hong Kong.

References

14. Ying XH, Zha HB, Identical projective geometric properties of central catadioptric line images and sphere images with applications to calibration, IJCV, 78, 89-105,(2008)