

Geometric Invariants and Applications under Catadioptric Camera Model

Yihong Wu and Zhanyi Hu

*National Laboratory of Pattern Recognition, Institute of Automation
Chinese Academy of Sciences, P.O.Box 2728, Beijing 100080, P. R. China
{yhwu, huzzy}@nlpr.ia.ac.cn*

Abstract

This paper presents geometric invariants of points and their applications under central catadioptric camera model. Although the image has severe distortions under the model, we establish some accurate projective geometric invariants of scene points and their image points. These invariants, being functions of principal point, are useful, from which a method for calibrating the camera principal point and a method for recovering planar scene structures are proposed. The main advantage of using these invariants for plane reconstruction is that neither camera motion nor the intrinsic parameters, except for the principal point, is needed. The theoretical correctness of the established invariants and robustness of the proposed methods are demonstrated by experiments. In addition, our results are found to be applicable to some more general camera models other than the catadioptric one.

1. Introduction

Catadioptric cameras are devices of conventional cameras combined with mirrors. As they have a large field of view, the catadioptric cameras have been widely used in surveillance, robot navigation and visualization.

Since Nayar [15] reported the catadioptric camera systems with a single viewpoint constraint in 1997, more and more attentions have been paid to this kind of omnidirectional camera models. Many geometric properties and methods for camera calibration and 3D reconstruction under the catadioptric models have been reported [1-10, 13-15, 17-20]. Geyer and Daniilidis [8] proposed a unified model for the devices with parabolic, elliptic, hyperbolic, and planar mirrors, which is equivalent to a two-step mapping through a sphere. This unified model can let us develop systematic theories for all these devices together conveniently.

Notice that in the past the entity of lines [3, 4, 7, 8, 9, 10, 20] has attracted more attentions than the entity of points [1, 13, 14] under the catadioptric cameras. We think further studies on points are needed because point is the most basic geometric feature. We will focus on investigating some geometric invariants of points in this paper. This is also motivated by the importance of geometric invariants in computer vision, which has been extensively shown under the perspective camera. However, due to the severe distortions of the view and the nonlinearity of the imaging process under the catadioptric cameras, the usual projective geometric properties are not preserved and the well-established theories under the perspective camera model cannot be applied.

The main contributions of this paper are two-fold:

1): The invariant equations for 1D, 2D and 3D space points are established via cross ratios of space points and the optical axis. These equations are invariant under the projective transformations of world coordinate system and image coordinate system. Although our invariant equations are derived from the central catadioptric model, they are in fact valid for a more general model as: The locus of optical centers lies on a segment of the optical axis rather than a single viewpoint and with only radial distortion.

2): The established invariant equations are applied to calibrating the camera principal point from 1D space points and to reconstructing a plane from at least five points and found to be effective. The proposed method for a plane reconstruction is an extension of the triangulation under a perspective camera [11].

Before starting our main work, we would like to point out: The established invariant equations vanish under the perspective model without distortion. This is because under the perspective camera, these equations become identities if the optical axis is replaced with an arbitrary space line. It follows that the principal point is not the unique point subject to the constraints and so cannot be determined.

The organization of the paper is as follows: Next section lists some preliminaries; the invariant equations are given in Section 3. Section 4 reports the invariant-based principal point calibration and plane reconstruction. Experiments are performed in Section 5, and Section 6 is some conclusions.

2. Preliminaries

A bold number or bold letter denotes a vector or homogeneous coordinates.

A pencil of lines is a set of lines through a fixed point. The fixed point is called the vertex of the pencil. For a pencil of four coplanar lines, a projective geometric invariant (a cross ratio) can be computed from 1D line-coordinates of these lines. This cross ratio is also equal to the cross ratio of the four intersecting points of a line with the four lines of this pencil [16].

A pencil of planes is a set of planes in space through a fixed line, and the fixed line is called the axis of the pencil. Similarly, there is a cross ratio from a pencil of four planes. This cross ratio is equal to the cross ratio of four intersecting lines of a space plane with the planes of the pencil. It is also equal to the cross ratio of the four intersecting points of a space line with the four planes of the pencil. [16].

We use the following notations:

$[]$ (a bracket):	the determinant of vectors in it
\mathbf{AB} (not in a bracket):	the line through \mathbf{A} and \mathbf{B}
$\mathbf{A}(\mathbf{C}, \mathbf{D}, \mathbf{E}, \mathbf{F})$:	the pencil of lines with \mathbf{A} as the vertex and through \mathbf{C} , \mathbf{D} , \mathbf{E} , \mathbf{F}
$\mathbf{AB}(\mathbf{C}, \mathbf{D}, \mathbf{E}, \mathbf{F})$:	the pencil of planes with \mathbf{AB} as the axis and through \mathbf{C} , \mathbf{D} , \mathbf{E} , \mathbf{F}

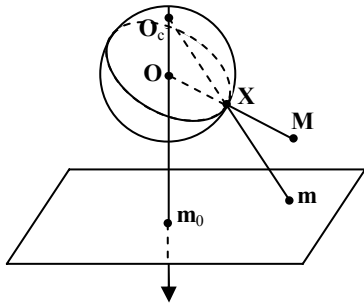


Figure 1. The imaging process under the central catadioptric model

The central catadioptric model is an equivalent spherical projection defined by Geyer and Daniilidis [8] as shown in Figure 1. We briefly recall this model here. A space point \mathbf{M} is projected to a point \mathbf{X} on

the viewing sphere through the sphere center \mathbf{O} , and then projected to \mathbf{m} on the image plane through the camera viewpoint \mathbf{O}_c . Let l be the distance from \mathbf{O} to \mathbf{O}_c , then the mathematical equations describing this imaging process involve camera intrinsic parameters, camera motion parameters, and l . The camera optical axis is the line \mathbf{OO}_c . The mirror used in this model is a paraboloid if $l=1$, an ellipsoid or hyperboloid if $0 < l < 1$, and a plane if $l=0$. The details can be found in [8].

3. Geometric invariants of scene points and image points

Let $\mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4}$ be four points in space, $\mathbf{m}_i, i=1..4$ be their image points, \mathbf{m}_0 the principal point of the camera, i.e. the intersection of the optical axis \mathbf{OO}_c with the image plane.

3.1. Invariant equation from 1D scene points

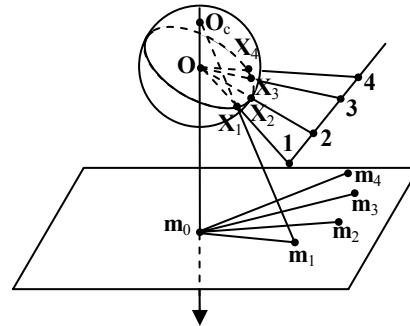


Figure 2. Equality of the cross ratio of $\mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4}$ and the cross ratio of $\mathbf{m}_0(\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3, \mathbf{m}_4)$

If $\mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4}$ are collinear, and without confusion, they are also used for their 1D homogeneous coordinates, then there is:

$$\frac{[\mathbf{13}][\mathbf{24}]}{[\mathbf{23}][\mathbf{14}]} = \frac{[\mathbf{m}_1\mathbf{m}_3\mathbf{m}_0][\mathbf{m}_2\mathbf{m}_4\mathbf{m}_0]}{[\mathbf{m}_2\mathbf{m}_3\mathbf{m}_0][\mathbf{m}_1\mathbf{m}_4\mathbf{m}_0]}. \quad (1)$$

The left side of (1) is the cross ratio of $\mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4}$, the right side of (1) is the cross ratio of the pencil of lines $\mathbf{m}_0(\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3, \mathbf{m}_4)$. The equation (1) means that the two cross ratios are equal. This is because: As shown in Figure 2, each space point \mathbf{i} and its image point \mathbf{m}_i are coplanar with the optical axis \mathbf{OO}_c , therefore, the intersections of the pencil of planes $\mathbf{OO}_c(\mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4})$ with the image plane is the pencil of lines $\mathbf{m}_0(\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3, \mathbf{m}_4)$, so the cross ratio of $\mathbf{m}_0(\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3, \mathbf{m}_4)$ is equal to the cross ratio of $\mathbf{OO}_c(\mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4})$. Moreover, the cross ratio of $\mathbf{OO}_c(\mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4})$ is equal to the cross ratio of $\mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4}$ since $\mathbf{1}, \mathbf{2}, \mathbf{3}, \mathbf{4}$ are collinear. Thus, (1) is true.

Consider the correspondences $\mathbf{m}'_i, i=1..4$ of $\mathbf{m}_i, i=1..4$ in another view, we also have:

$$\frac{[13][24]}{[23][14]} = \frac{[\mathbf{m}'_1 \mathbf{m}'_3 \mathbf{m}'_0][\mathbf{m}'_2 \mathbf{m}'_4 \mathbf{m}'_0]}{[\mathbf{m}'_2 \mathbf{m}'_3 \mathbf{m}'_0][\mathbf{m}'_1 \mathbf{m}'_4 \mathbf{m}'_0]},$$

where \mathbf{m}'_0 is the principal point of the second view. Therefore, the following invariant equation holds from two views:

$$\frac{[\mathbf{m}_1 \mathbf{m}_3 \mathbf{m}_0][\mathbf{m}_2 \mathbf{m}_4 \mathbf{m}_0]}{[\mathbf{m}_2 \mathbf{m}_3 \mathbf{m}_0][\mathbf{m}_1 \mathbf{m}_4 \mathbf{m}_0]} = \frac{[\mathbf{m}'_1 \mathbf{m}'_3 \mathbf{m}'_0][\mathbf{m}'_2 \mathbf{m}'_4 \mathbf{m}'_0]}{[\mathbf{m}'_2 \mathbf{m}'_3 \mathbf{m}'_0][\mathbf{m}'_1 \mathbf{m}'_4 \mathbf{m}'_0]}. \quad (2)$$

This means that there is a homography between $\mathbf{m}_0(\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3, \mathbf{m}_4)$ and $\mathbf{m}'_0(\mathbf{m}'_1, \mathbf{m}'_2, \mathbf{m}'_3, \mathbf{m}'_4)$.

The equations (1) and (2) can be extended to coplanar scene lines by considering the intersection points of these lines and considering the images of the intersection points.

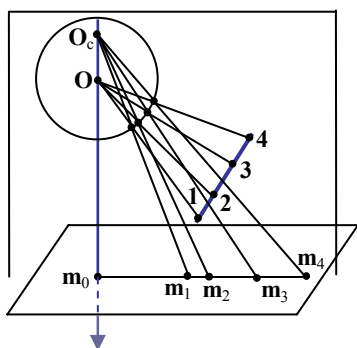


Figure 3. \mathbf{m}_0 and all \mathbf{m}_i are collinear when the line through 1, 2, 3, 4 are coplanar with OO_c .

Degeneracy: If the line passing through 1, 2, 3, 4 is coplanar with the camera optical axis OO_c , the images $\mathbf{m}_i, i=1..4$ will be collinear in the image plane, as shown in Figure 3. At the time, the pencil of planes $OO_c(1, 2, 3, 4)$ becomes coincident, (1) no longer holds. But, we still have another constraint on \mathbf{m}_0 , i.e., \mathbf{m}_0 and all $\mathbf{m}_i, i=1..4$ are collinear. It should be noted that although $\mathbf{m}_i, i=1..4$ are collinear and 1, 2, 3, 4 are collinear, the cross ratio of $\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3, \mathbf{m}_4$ is in general not equal to the cross ratio of 1, 2, 3, 4. The two cross ratios are equal if and only if $O, O_c, X_1, X_2, X_3, X_4$ are con-conic. Because by Chasles's Theorem [16] the necessary and sufficient condition that $O, O_c, X_1, X_2, X_3, X_4$ are con-conic is that the cross ratio of $O(X_1, X_2, X_3, X_4)$ is equal to the cross ratio of $O_c(X_1, X_2, X_3, X_4)$, where the cross ratio of $O(X_1, X_2, X_3, X_4)$ is equal to the cross ratio of 1, 2, 3, 4, and the cross ratio of $O_c(X_1, X_2, X_3, X_4)$ is equal to the cross ratio of $\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3, \mathbf{m}_4$.

3.2. Invariant equation from 2D scene points

If 1, 2, 3, 4 are noncollinear, the invariant equation (1) no longer holds. But, we still have equations on some invariants of scene points and their image points.

We consider the case that space points 1, 2, 3, 4, also used as their 2D homogeneous coordinates, are coplanar. The base plane of these points is denoted as π , and the intersection point of π with the optical axis OO_c is denoted as \mathbf{M}_0 , as shown in Figure 4. Then the invariant equation can be established via \mathbf{M}_0 as:

$$\frac{[13\mathbf{M}_0][24\mathbf{M}_0]}{[23\mathbf{M}_0][14\mathbf{M}_0]} = \frac{[\mathbf{m}_1 \mathbf{m}_3 \mathbf{m}_0][\mathbf{m}_2 \mathbf{m}_4 \mathbf{m}_0]}{[\mathbf{m}_2 \mathbf{m}_3 \mathbf{m}_0][\mathbf{m}_1 \mathbf{m}_4 \mathbf{m}_0]}. \quad (3)$$

This equation means that the cross ratio of the pencil of lines $\mathbf{M}_0(1, 2, 3, 4)$ is equal to the cross ratio of the pencil of lines $\mathbf{m}_0(\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3, \mathbf{m}_4)$. The reason is as follows. As shown in Figure 4, the space plane π intersects with the pencil of planes $OO_c(1, 2, 3, 4)$ at the pencil of lines $\mathbf{M}_0(1, 2, 3, 4)$, and the image plane intersects with the pencil of planes $OO_c(1, 2, 3, 4)$ at the pencil of lines $\mathbf{m}_0(\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3, \mathbf{m}_4)$. Therefore, the two cross ratios of $\mathbf{M}_0(1, 2, 3, 4)$ and $\mathbf{m}_0(\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3, \mathbf{m}_4)$ are both equal to the cross ratio of the pencil of planes $OO_c(1, 2, 3, 4)$. So, (3) is true.

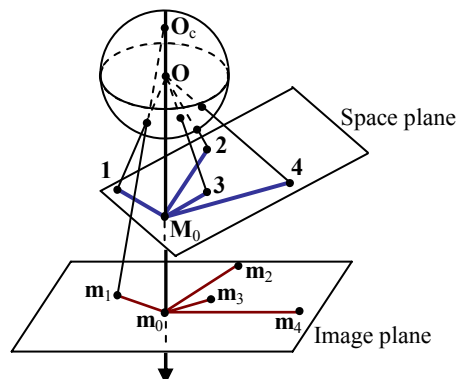


Figure 4. Equality of the two cross ratios of $\mathbf{M}_0(1, 2, 3, 4)$ and $\mathbf{m}_0(\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3, \mathbf{m}_4)$

Let \mathbf{M}'_0 be the intersection point of the optical axis of the second camera with π . If $\mathbf{M}_0 = \mathbf{M}'_0$, (2) still holds in the 2D case. Otherwise, (2) no longer holds in the 2D case.

Degeneracy: If the plane containing 1, 2, 3, 4 is coplanar with the camera optical axis OO_c , then the images $\mathbf{m}_i, i=1..4$ will be collinear in the image plane, see Figure 5. At the time, the pencil of planes $OO_c(1, 2, 3, 4)$ becomes coincident, (3) no longer holds. But, we still have another constraint on \mathbf{m}_0 , i.e., \mathbf{m}_0 and all $\mathbf{m}_i, i=1..4$ are collinear as shown in Figure 5.

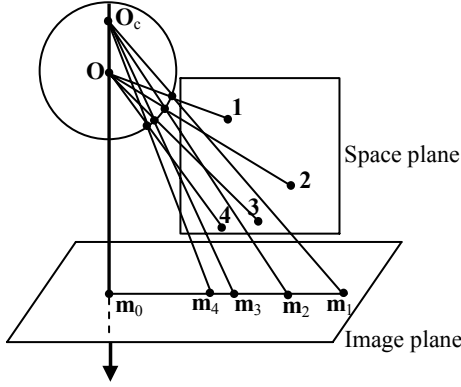


Figure 5. \mathbf{m}_0 and all $\mathbf{m}_i, i=1..4$ are collinear when the plane containing 1, 2, 3, 4 are coplanar with \mathbf{OO}_c

3.3. Invariant equation from 3D scene points

If 1, 2, 3, 4 are neither collinear nor coplanar, (1), (2) and (3) no longer hold. The invariant equation for the 3D scene points is:

$$\frac{[\mathbf{OO}_c \mathbf{13}][\mathbf{OO}_c \mathbf{24}]}{[\mathbf{OO}_c \mathbf{23}][\mathbf{OO}_c \mathbf{14}]} = \frac{[\mathbf{m}_1 \mathbf{m}_3 \mathbf{m}_0][\mathbf{m}_2 \mathbf{m}_4 \mathbf{m}_0]}{[\mathbf{m}_2 \mathbf{m}_3 \mathbf{m}_0][\mathbf{m}_1 \mathbf{m}_4 \mathbf{m}_0]}, \quad (4)$$

where 1, 2, 3, 4 are 3D homogeneous coordinates, the left side is the cross ratio of $\mathbf{OO}_c(1, 2, 3, 4)$, and the right side is the cross ratio of $\mathbf{m}_0(\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3, \mathbf{m}_4)$. (4) is useful for 3D reconstruction by solving the optical axis \mathbf{OO}_c and for object recognition by eliminating the optical axis \mathbf{OO}_c . Due to the space limit, further discussions on the 3D case are omitted.

Extension of the invariant equations to more general camera models: The derivations of the above invariant equations (1), (2), (3), (4) and the degenerate discussions are based on the camera model defined by Geyer and Daniilidis [8], where camera optical center \mathbf{O}_c is a single point. In fact, the condition that \mathbf{O}_c is a single point is not strictly required. If the locus of \mathbf{O}_c is a line segment on the optical axis and image points have only radial distortion, the above invariant equations still hold because for such a model, the required condition that $\mathbf{OO}_c, \mathbf{m}_i$ and its corresponding space point \mathbf{i} are coplanar is also satisfied. Thus, perspective camera with only radial distortion or some kind of fisheye cameras can use the results of this paper directly.

Under the perspective camera, if the view has no distortion, the above invariant equations vanish. This is because if an arbitrary space line substitutes for \mathbf{OO}_c and if the intersection points of this space line with the image plane and with the scene plane substitutes for

$\mathbf{m}_0, \mathbf{M}_0$, these equations are still holds. Therefore, $(\mathbf{OO}_c, \mathbf{m}_0, \mathbf{M}_0)$ is not the unique entity subject to the constraints.

4. Calibration and reconstruction from geometric invariants of points

The above-established invariant equations in Section 3 find their applications in camera calibration, image matching, 3D reconstruction, and object recognition under catadioptric cameras. Two applications are presented below. The advantages of our methods on calibration and reconstruction are at least two-fold: Firstly, it is simple using the geometric invariants; secondly, we can recover directly the scene structure without prior camera pose determination and distortion correction.

4.1. Calibrating principal point from 1D objects

Let S be a set of quadruplets of collinear scene points. The steps to calibrate the principal point \mathbf{m}_0 from these quadruplets and their images are:

Take each quadruplet from S , and denote its elements as 1, 2, 3, 4, and the corresponding image points as $\mathbf{m}_i, i=1..4$.

Step 1. Set up invariant equations on \mathbf{m}_0 .

Case (i): When $\mathbf{m}_i, i=1..4$ are noncollinear and the structure of 1, 2, 3, 4 is known

From a single view, we have the following quadric constraint on \mathbf{m}_0 by (1):

$$[\mathbf{23}][\mathbf{14}][\mathbf{m}_1 \mathbf{m}_3 \mathbf{m}_0][\mathbf{m}_2 \mathbf{m}_4 \mathbf{m}_0] - [\mathbf{13}][\mathbf{24}][\mathbf{m}_2 \mathbf{m}_3 \mathbf{m}_0][\mathbf{m}_1 \mathbf{m}_4 \mathbf{m}_0] = 0,$$

where \mathbf{m}_0 is unknown with two elements.

Case (ii): When $\mathbf{m}_i, i=1..4$ are noncollinear and the structure of 1, 2, 3, 4 is unknown

From two views, we have the following constraint on the principal points $\mathbf{m}_0, \mathbf{m}'_0$ as:

$$[\mathbf{m}_1 \mathbf{m}_3 \mathbf{m}_0][\mathbf{m}_2 \mathbf{m}_4 \mathbf{m}_0][\mathbf{m}'_2 \mathbf{m}'_3 \mathbf{m}'_0][\mathbf{m}'_1 \mathbf{m}'_4 \mathbf{m}'_0] - [\mathbf{m}_2 \mathbf{m}_3 \mathbf{m}_0][\mathbf{m}_1 \mathbf{m}_4 \mathbf{m}_0][\mathbf{m}'_1 \mathbf{m}'_3 \mathbf{m}'_0][\mathbf{m}'_2 \mathbf{m}'_4 \mathbf{m}'_0] = 0$$

If $\mathbf{m}_0 \neq \mathbf{m}'_0$, \mathbf{m}_0 and \mathbf{m}'_0 can be determined from at least four such constraints. We consider $\mathbf{m}_0 = \mathbf{m}'_0$ in the next.

Case (iii): When $\mathbf{m}_i, i=1..4$ are collinear

At the time, the structure of 1, 2, 3, 4 is not required to be known, and $\mathbf{m}_i, i=1..4$ are collinear with \mathbf{m}_0 . From a single view, we set up a constraint on \mathbf{m}_0 as:

$$\mathbf{l}^T \mathbf{m}_0 = 0,$$

where \mathbf{l} is the extracted line-coordinates of the line through $\mathbf{m}_i, i=1..4$.

By repeating the above process, we set up all invariant equations on \mathbf{m}_0 from all quadruplets in S . In order to determine \mathbf{m}_0 uniquely, at least two quadruplets are required.

Step 2. Solve the equations in Step 1 for $\mathbf{m}_0 = (x \ y \ 1)^T$ and take the result as the initial value of \mathbf{m}_0 . Two cases are discussed below.

Case (a): If there are just two quadruplets in S , we obtain two equations on \mathbf{m}_0 . Solve the equations and choose the real solution close to the center of the view as the initial value of \mathbf{m}_0 .

Case (b): If more than two linear equations on \mathbf{m}_0 are available, SVD approach is used. If more than two quadric equations on \mathbf{m}_0 are available, we construct a matrix, whose each row is the coefficients of x^2 , xy , y^2 , x , y , 1 of each equation, then a more robust SVD-like approach is used. Otherwise, from the equations Groebner basis method is used to obtain lower order univariate equations and then SVD-like approach is engaged. The details are omitted due to the space limit.

Step 3. Do post-optimization for \mathbf{m}_0 by the Levenberg-Marquadt algorithm. The cost function is the sum of the residual squares in the above equations of Step 1 for all possible quadruplets.

4.2. Direct reconstruction of plane structure

Let π be a scene plane noncoplanar with camera optical axes, and $1, 2, 3, 4, 5$ five known points on π , also used to denote their 2D homogeneous coordinates. Then without correcting distortions and calibrating other camera parameters except for the principal point, we can recover the whole structure of π from two views. The process is as follows and is shown in Figure 6, where \mathbf{X} is a point of π to be reconstructed.

Step 1. Calibrate the principal points $\mathbf{m}_0, \mathbf{m}'_0$ in the two views, for example by the method in Section 4.1, or those of [3, 7, 9].

Step 2. Find out the intersection points $\mathbf{M}_0, \mathbf{M}'_0$ of the two camera optical axes with the scene plane π .

Choose four points $\mathbf{i}, \mathbf{j}, \mathbf{k}, \mathbf{h}$ from $1, 2, 3, 4, 5$, then from (3), we have:

$$\frac{[\mathbf{i} \ \mathbf{k} \ \mathbf{M}_0][\mathbf{j} \ \mathbf{h} \ \mathbf{M}_0]}{[\mathbf{m}_i \ \mathbf{m}_k \ \mathbf{m}_0][\mathbf{m}_j \ \mathbf{m}_h \ \mathbf{m}_0]} [\mathbf{i} \ \mathbf{h} \ \mathbf{M}_0][\mathbf{j} \ \mathbf{k} \ \mathbf{M}_0] = 0,$$

where only \mathbf{M}_0 is unknown. We set up all such equations from four points out of $1, 2, 3, 4, 5$. Among these equations, only two are independent, from which \mathbf{M}_0 can be determined uniquely and linearly by polynomial elimination and factorization. The detail is

omitted due to the space limit. Similarly, \mathbf{M}'_0 can be determined.

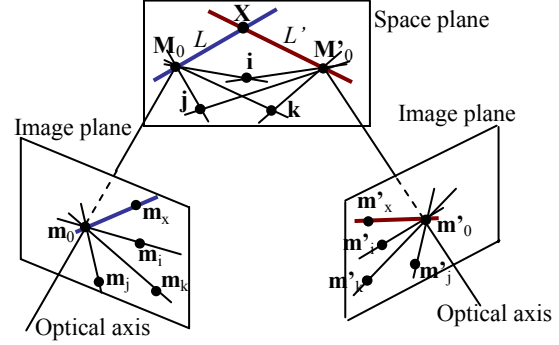


Figure 6. The process to recover the structure of a plane

Step 3. Let $\mathbf{m}_x, \mathbf{m}'_x$ be the corresponding image points in the two views of the scene point \mathbf{X} . Reconstruct the two lines $\mathbf{m}_0\mathbf{m}_x, \mathbf{m}'_0\mathbf{m}'_x$ on π , the results are denoted as L, L' , see Figure 6.

For \mathbf{X} and for $\mathbf{i}, \mathbf{j}, \mathbf{k}$ out of $1, 2, 3, 4, 5$, from (3) there is:

$$\frac{[\mathbf{i} \ \mathbf{k} \ \mathbf{M}_0][\mathbf{j} \ \mathbf{X} \ \mathbf{M}_0]}{[\mathbf{m}_i \ \mathbf{m}_k \ \mathbf{m}_0][\mathbf{m}_j \ \mathbf{m}_x \ \mathbf{m}_0]} [\mathbf{j} \ \mathbf{k} \ \mathbf{M}_0][\mathbf{i} \ \mathbf{X} \ \mathbf{M}_0] = 0, \quad (5)$$

where \mathbf{M}_0 is given in Step 2 and only \mathbf{X} is unknown. The locus of \mathbf{X} satisfying (5) is a line, the reconstruction \mathbf{M}_0 of \mathbf{m}_0 is on this line and the reconstruction \mathbf{X} of \mathbf{m}_x is also on this line. Thus, this line is the reconstruction of the line $\mathbf{m}_0\mathbf{m}_x$, i.e. it is the equation of L . Different $\mathbf{i}, \mathbf{j}, \mathbf{k}$ out of $1, 2, 3, 4, 5$ ideally should give the same equation of L . However, all such equations are used for the robustness purpose in this work.

Similarly, from the second view, we obtain:

$$\frac{[\mathbf{i} \ \mathbf{k} \ \mathbf{M}'_0][\mathbf{j} \ \mathbf{X} \ \mathbf{M}'_0]}{[\mathbf{m}'_i \ \mathbf{m}'_k \ \mathbf{m}'_0][\mathbf{m}'_j \ \mathbf{m}'_x \ \mathbf{m}'_0]} [\mathbf{j} \ \mathbf{k} \ \mathbf{M}'_0][\mathbf{i} \ \mathbf{X} \ \mathbf{M}'_0] = 0,$$

It is the equation of L' through \mathbf{M}'_0 and \mathbf{X} .

Step 4. The scene point \mathbf{X} of \mathbf{m}_x and \mathbf{m}'_x on π satisfies all the equations in Step 3. These equations are all linear on \mathbf{X} . SVD method is used to obtain the solution of \mathbf{X} .

Step 5. Levenberg-Marquardt optimization is followed. The cost function is the sum of the residual squares of the invariant equations in Step 3.

In the above method, at least five scene points from two views should be known. More scene points and more views can improve the robustness and accuracy.

Degeneracy: If the two optical axes intersect π at the same point, i.e., $\mathbf{M}_0 = \mathbf{M}'_0$, then the reconstructions

L and L' are identical, so the reconstruction \mathbf{X} cannot be determined uniquely, it has one degree of freedom on the line L . Another similar degenerate case is when the scene point lies on the line through $\mathbf{M}_0\mathbf{M}'_0$, L and L' become coincident. More views should be used to remove these ambiguities.

Our method to reconstruct \mathbf{X} is an extension of the triangulation method for perspective projections [11]. By back projecting $\mathbf{m}_0\mathbf{m}_x$, $\mathbf{m}'_0\mathbf{m}'_x$, we obtain L and L' on the scene plane. Their unique intersection point is the reconstruction of \mathbf{X} . Not only a plane structure but also 3D scene points can be reconstructed. This method to reconstruct a 3D scene point from three catadioptric views is as follows: Let \mathbf{m}''_0 be the principal point and \mathbf{m}''_x be the image point of \mathbf{X} in the third view. Then, three space planes, each containing a different optical axis, can be obtained by back projecting of $\mathbf{m}_0\mathbf{m}_x$, $\mathbf{m}'_0\mathbf{m}'_x$, $\mathbf{m}''_0\mathbf{m}''_x$. The unique intersection point of these three planes is the reconstruction of \mathbf{X} . This work will be reported in future.

5. Experiments

In this section, experiments are carried out to test the proposed invariant equations in Section 3 and application algorithms in Section 4. We use a perspective camera with a hyperbolic mirror.

5.1. Simulations

The simulated camera has the following parameters $f_u = 610$, $f_v = 600$, $s = 0.8$, $u_0 = 500$, $v_0 = 350$, $l = 0.9231$, where f_u , f_v are the focal lengths, s is the skew factor, (u_0, v_0) is the principal point, l is the distance from \mathbf{O} to \mathbf{O}_c . The parameters keep unchanged when camera is moving. We take 16 space points on $X-Y$ plane shown in (a) of Figure 7, then project them to the catadioptric image planes at three different positions, see (b), (c), (d) of Figure 7 for the three images. The image sizes are not greater than 1000×1000 pixels. The Gaussian noise with mean 0 and standard deviation ranging from 0 to 2.0 pixels is directly added to each of the image points, and then the principal point \mathbf{m}_0 and the intersection points \mathbf{M}_0 , \mathbf{M}'_0 , \mathbf{M}''_0 of the three optical axes with $X-Y$ plane are estimated by the algorithm in Subsection 4.1 and 4.2. At each noise level, we perform 100 trials, and the averaged results are shown in Table 1 and Table 2, where the values under noise level 0 are the same as their ground truths. From these results, we can see the stability and validity of our method. The variances of

the estimated \mathbf{m}_0 under different noise levels are also computed and plotted in Figure 8. One of the variances from all the estimations of \mathbf{M}_0 , \mathbf{M}'_0 and \mathbf{M}''_0 under all noise levels is 1.0908, and all of the others are smaller than 1.

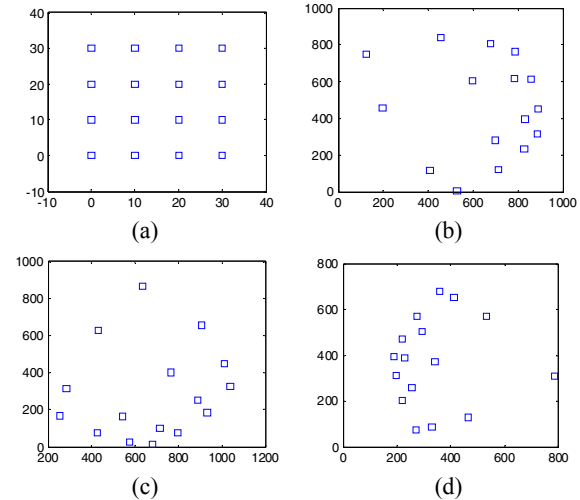


Figure 7. (a): the scene points; (b), (c) and (d): the three views of (a)

Table 1. The averages of the estimated principal point vs. noise

Noise level (Pixel)	Principal point
0	(500.00, 350.00)
0.4	(500.50, 350.21)
0.8	(500.28, 350.21)
1.2	(500.36, 349.82)
1.6	(502.48, 350.42)
2.0	(505.16, 349.06)

Table 2. The averages of the estimated intersection points of the three optical axes with the $X-Y$ plane vs. noise (pixel)

Noise Level	\mathbf{M}_0	\mathbf{M}'_0	\mathbf{M}''_0
0	(5.00, 15.00)	(15.00, 25.00)	(25.00, 5.00)
0.4	(5.02, 15.00)	(15.02, 24.99)	(25.03, 5.01)
0.8	(4.98, 14.98)	(15.03, 25.02)	(24.97, 4.97)
1.2	(4.99, 14.98)	(14.93, 24.95)	(24.97, 4.95)
1.6	(5.07, 14.99)	(14.98, 24.92)	(25.07, 4.98)
2	(5.22, 15.00)	(15.20, 24.99)	(25.16, 4.92)

The above space points and the estimations of \mathbf{m}_0 , \mathbf{M}_0 , \mathbf{M}'_0 , \mathbf{M}''_0 are used to reconstruct other space points of the $X-Y$ plane shown in Figure 9. From the contaminated images with Gaussian noise as described before, the algorithm of Subsection 4.2 is used, and the averaged result of 100 independent experiments is computed. Out of all these estimations, the biggest

error is 0.7145. The variances are also calculated showing that most of them are smaller than 1 and the biggest one is 2.0587. The estimations of two space points with bigger errors are shown in Table 3 where the values under noise level 0 are the same as their ground truths.

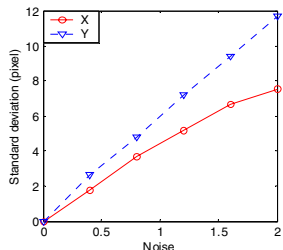


Figure 8. The std. of the estimated principal point vs. noise

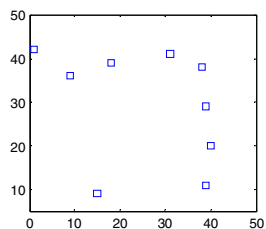


Figure 9. The scene points to be reconstructed

Table 3. Reconstructions of two scene points from three views

Noise Level (Pixel)	Reconstructions	
0	(1.00, 42.00)	(18.00, 39.00)
0.4	(0.99, 42.02)	(17.99, 39.00)
0.8	(0.87, 42.25)	(18.05, 39.21)
1.2	(0.96, 41.82)	(17.89, 38.95)
1.6	(0.91, 41.89)	(17.85, 38.94)
2.0	(0.99, 42.69)	(18.25, 39.71)

5.2. Experiments with real data

In the following, the results in Section 3 and Section 4 are tested by real data.

We conduct experiments from two real images. The images are of size 2048×1536 pixels captured by a NIKON COOLPIX990 camera with a hyperbolic mirror designed by the Center for Machine Perception, Czech Technical University, as shown in Figure 10. The image point extraction and matching are currently done manually. The geometric information of 1D scene points on the ceiling is used to calibrate each principal point from each single view, and the results are:

$$\mathbf{m}_0 = (1028.68, 777.55, 1)^T, \mathbf{m}'_0 = (1062.36, 773.23, 1)^T$$

In order to assess the accuracy of the estimations, some cross ratios from image points and the estimated \mathbf{m}_0 in

the first view are computed. The results are shown in Table 4. It can be seen that the estimated invariants are quite close to their ground truths. We also found that the estimated cross ratios from the second view and the estimated \mathbf{m}'_0 are quite close to their ground truths.



Figure 10. The used real images

Table 4. The estimations of some cross ratios and their ground truths (GT)

Estimation	GT	Estimation	GT
1.3274	1.3333	0.3325	0.3333
0.4439	0.4444	0.4988	0.5
4.0188	4	-0.4874	-0.5
-0.7777	-0.8	-0.9727	-1
0.2574	0.25	-2.0488	-2
-1.2770	-1.25	-1.0249	-1

To uniquely determine the solutions of the reconstruction by our method, the minimal number of the known scene points is five. Here, we are with the minimal number of the known scene points. Let the ceiling be the X - Y plane of the world coordinate system. We use the scene points $\mathbf{M}_1 = (0, 0, 1)$, $\mathbf{M}_2 = (1, 0, 1)$, $\mathbf{M}_3 = (0, 1, 1)$, $\mathbf{M}_4 = (1, 1, 1)$, $\mathbf{M}_5 = (2, 0, 1)$, whose image points are \mathbf{m}_i , $i=1..5$ show in the left figure of Figure 10, to recover the coordinates of other scene points on the ceiling.

The estimated intersection points of the two optical axes with the ceiling are respectively:

$$\mathbf{M}_0 = (-4.68, -5.27, 1)^T, \mathbf{M}'_0 = (-0.48, -7.01, 1)^T.$$

Then, the equations on other unknown scene point are set up, and the coordinates are solved. The results are shown in Table 5. They demonstrate the validity of the invariant equations in Section 3 as well as the method for recovering the structure of a plane in Section 4.

Table 5. The estimated coordinates of some scene points on the ceiling and their ground truths (GT)

Estimation	GT	Estimation	GT
(-3.05, 2.10)	(-3, 2)	(-2.04, 2.02)	(-2, 2)
(-0.02, 2.01)	(0, 2)	(0.99, 2.00)	(1, 2)
(-2.99, 1.11)	(-3, 1)	(-2.00, 1.05)	(-2, 1)
(2.02, 1.02)	(2, 1)	(3.05, 1.06)	(3, 1)
(-1.98, 0.07)	(-2, 0)	(-0.99, 0.04)	(-1, 0)
(-2.95, 0.76)	(-3, 1)	(-1.95, -0.88)	(-2, -1)
(-0.01, 0.99)	(0, 1)	(0.97, -1.00)	(1, -1)
(2.99, -0.96)	(3, -1)	(-0.97, -1.94)	(-1, -2)
(0.92, -1.99)	(1, -2)	(1.91, -1.98)	(2, -2)
(-1.02, 2.01)	(-1, 2)	(2.01, 2.01)	(2, 2)
(-1.01, 0.99)	(-1, 1)	(-2.96, 0.17)	(-3, 0)
(3.03, 0.05)	(3, 0)	(-0.97, -0.92)	(-1, -1)
(1.97, -0.99)	(2, -1)	(-0.04, -1.98)	(0, -2)

6. Conclusions

Three kinds of interesting invariant equations from 1D, 2D and 3D scene points are introduced under the catadioptric camera model. Their applications for calibrating the principal point or reconstructing plane structure are illustrated. The experimental results confirm the theoretical correctness of our invariant equations and show the accuracy and stability of our methods. These invariants are also useful for object recognition and image matching. Their further applications will be investigated.

Acknowledgements

This work was supported by the National Key Basic Research and Development Program (973) under grant No. 2002CB312104, and the National Natural Science Foundation of China under grant No. 60303021.

References

[1] D.G. Aliaga, "Accurate catadioptric calibration for real-time pose estimation in room-size environments", *The 8th International Conference on Computer Vision*, 2001, pp. 127-134.

[2] J. Baldwin and A. Basu, "3D estimation using panoramic stereo", *The 15th International Conference on Pattern Recognition*, 2000, pp. 97-100.

[3] J.P. Barreto and H. Araujo, "Geometric properties of central catadioptric line images", *The 7th European Conference on Computer Vision*, 2002, pp. 237-251.

[4] J.P. Barreto and H. Araujo, "Direct least square fitting of paracatadioptric line images", *CVPR2003 Workshop on Omnidirectional Vision and Camera Networks*, 2003, pp. 78-83.

[5] E.L.L. Cabral, J.C. de Souza, and M.C. Hunold, "Omnidirectional stereo vision with a hyperbolic

double lobed mirror", *The 17th International Conference on Pattern Recognition*, 2004, pp. 1-4.

[6] P. Doubek and T. Svoboda, "Reliable 3D reconstruction from a few catadioptric images", *The 3rd Workshop on Omnidirectional Vision*, 2002, pp. 71-78.

[7] C. Geyer and K. Daniilidis, "Catadioptric camera calibration", *The 7th International Conference on Computer Vision*, 1999, pp. 398-403.

[8] C. Geyer and K. Daniilidis, "Catadioptric projective geometry", *International Journal of Computer Vision*, 2001, Vol. 45, No. 3, pp. 223-243.

[9] C. Geyer and K. Daniilidis, "Paracatadioptric camera calibration", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, Vol. 24, No. 5, pp. 687-695.

[10] C. Geyer and K. Daniilidis, "Properties of the catadioptric fundamental matrix", *The 7th European Conference on Computer Vision*, 2002, pp. 140-154.

[11] R. Hartley and P. Sturm, "Triangulation", *Computer Vision and Image Understanding*, Vol. 68, No. 2, pp. 146-157, 1997. Also: *6th International Conference on Computer Analysis of Images and Patterns*, 1995, pp. 190-197.

[12] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2000.

[13] S.B. Kang, "Catadioptric self-calibration", *IEEE Conference on Computer Vision and Pattern Recognition*, 2000, pp. 201-207.

[14] B. Micusik, T. Pajdla, "Para-catadioptric camera auto-calibration from epipolar Geometry", *The 6th Asian Conference on Computer Vision*, 2004, pp. 748-753.

[15] S.K. Nayar, "Catadioptric omnidirectional camera", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1997, pp. 482-488.

[16] J.G. Semple and G.T. Kneebone, *Algebraic projective geometry*, Oxford University Press, 1952.

[17] T. Svoboda, T. Pajdla, and V. Hlavac, "Epipolar geometry for panoramic cameras", *The 5th European Conference on Computer Vision*, 1998, pp. 218-231.

[18] T. Svoboda, T. Pajdla, and V. Hlavac, "Motion estimation using central panoramic cameras", *IEEE Conference on Intelligent Vehicles*, 1998, pp. 335-340.

[19] X.H. Ying and Z.Y. Hu, "Catadioptric camera calibration using geometric invariants", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2004, Vol. 26, No. 10, pp.1260-1271.

[20] X.H. Ying and Z.Y. Hu, "Catadioptric line features detection using Hough transform", *The 17th International Conference on Pattern Recognition*, 2004, pp. 839-842.