INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Generic Calibration of Axial Cameras

Srikumar Ramalingam — Peter Sturm — Suresh K. Lodha

N° 5827

December 2005

THÈME 3

ISSN 0249-6399 ISRN INRIA/RR--5827--FR+ENG

INRIA
RHÔNE-ALPES

Generic Calibration of Axial Cameras

Srikumar Ramalingam* , Peter Sturm† , Suresh K. Lodha‡

Thème 3 — Interaction homme-machine,
images, données, connaissances
Projet Movi

Rapport de recherche n° 5827 — December 2005 — 16 pages

Abstract: Although most works in computer vision use perspective or other central cameras, the interest in non-central camera models has increased lately, especially with respect to omnidirectional vision. Calibration and structure-from-motion algorithms exist for both, central and non-central cameras. An intermediate class of cameras, although encountered rather frequently, has received less attention. So-called *axial cameras* are non-central but their projection rays are constrained by the existence of a line that cuts all of them. This is the case for stereo systems, many non-central catadioptric cameras and pushbroom cameras for example. In this report, we study the geometry of axial cameras and propose a calibration approach for them. We also describe the various axial catadioptric configurations which are more common and less restrictive than central catadioptric ones. Finally we used simulations and real experiments to prove the validity of our theory.

Key-words: Calibration, camera model, distortion, stereo, omnidirectional, catadioptric

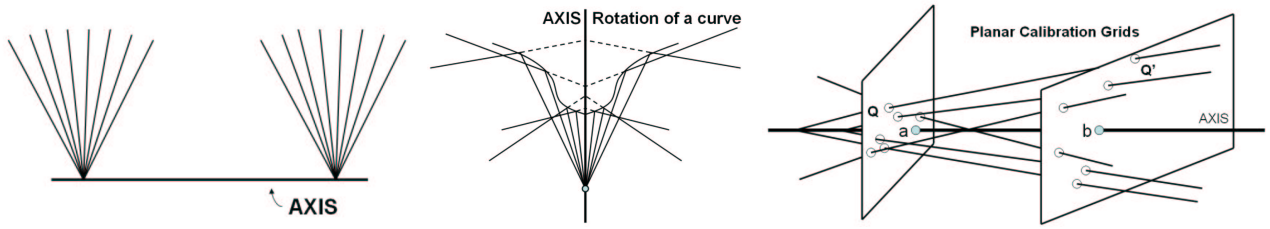
* srikumar@cse.ucsc.edu
† Peter.Sturm@inrialpes.fr
‡ lodha@cse.ucsc.edu

Unité de recherche INRIA Rhône-Alpes

Calibrage générique de caméras axiales

Résumé : La plupart des travaux en vision par ordinateur emploient le modèle de caméra perspective ou d'autres modèles de caméra centraux, c'est-à-dire à foyer unique. Récemment, l'intérêt pour des modèles non-centraux a augmenté considérablement, surtout en ce qui concerne la vision panoramique, ou bien omnidirectionnelle. Il existe des algorithmes de calibrage et de « structure-from-motion » pour des caméras centrales ou non-centrales. Une classe intermédiaire, bien que rencontrée assez fréquemment, a suscité moins d'attention : nous appelons *caméras axiales* celles dont les lignes de vue n'ont pas de foyer unique, mais où il existe une ligne qui les coupe toutes. C'est le cas par exemple pour les caméras de type push-broom, pour beaucoup de systèmes catadioptriques, ou encore pour chaque système stéréo, s'il est interprété comme un seul capteur. Dans ce rapport, nous étudions la géométrie des caméras axiales et en proposons une approche de calibrage. Nous décrivons également les diverses configurations catadioptriques axiales qui sont le plus communes et qui sont moins restrictives que les modèles catadioptriques centraux. Enfin, nous prouvons la validité de notre théorie par des expériences sur données simulées et réelles.

Mots-clés : Calibrage, modèle de caméra, distorsion, stéréo, omnidirectionnel, catadioptrique



(a) (b) (c)

Figure 1: Examples of axial imaging models (a) stereo camera (b) a mirror formed by rotating a planar curve about an axis containing the optical center of the perspective camera.(c) Calibration of axial cameras using calibration grids: The projection rays, camera axis and two grids are shown. The axis intersects at **a** and **b** on the first and the second calibration grids respectively.

1 Introduction

Many camera models have been considered in computer vision and related fields and even more taylor-made calibration methods have been developed. Most of those are designed for central cameras, but approaches and studies for non-central or general ones also exist [5, 8, 9, 10, 11, 13, 15, 17, 3]. An intermediate class of cameras, lying between central and fully non-central ones, is that of so-called *axial cameras*: their projection rays are constrained by the existence of a line that cuts all of them, the **camera axis**, but they may not go through a single optical center.

The axial model is a rather useful one (cf. figure 1(a) and (b)). Many misaligned catadioptric configurations fall under this model. Such configurations, which are slightly non-central, are usually classified as a non-central camera and calibrated using an iterative nonlinear algorithm [12, 1, 7]. For example, whenever the mirror is a surface of revolution and the central camera looking at the mirror lies anywhere on the revolution axis, the system is of axial type. Furthermore, two-camera stereo systems or systems consisting of three or more aligned cameras, are axial. Pushbroom cameras [6] are another example, although they are of a more restricted class (there exist two camera axes [4]).

In this report, we propose a generic calibration approach for axial cameras, the first to our knowledge. It uses images of planar calibration grids, put in unknown positions. We show the existence of multi-view tensors that can be estimated linearly and from which the pose of the calibration grids as well as the position of the camera axis, can be recovered. The actual calibration is then performed by computing projection rays for all individual pixels of a camera, constrained to cut the camera axis.

The report is organized as follows. The problem is formalized in section 2. In section 3, we show what can be done with two images of calibration grids. Complete calibration using three images, is described in section 4, followed by a bundle adjustment algorithm in section 6. Various types of axial catadioptric cameras are listed in section 7. Experimental results and conclusions are given in sections 8 and 9.

2 Problem Formulation

In the following, we will call **camera axis** the line cutting all projection rays. It will be represented by a 6-vector $\mathbf{L}$ and the associated 4×4 skew-symmetric Plücker matrix $[\mathbf{L}]_\times$:

$$[\mathbf{L}]_\times = \begin{pmatrix} 0 & -L_4 & L_6 & -L_2 \\ L_4 & 0 & -L_5 & -L_3 \\ -L_6 & L_5 & 0 & -L_1 \\ L_2 & L_3 & L_1 & 0 \end{pmatrix}$$

The product $[\mathbf{L}]_\times \mathbf{Q}$ gives the plane spanned by the line $\mathbf{L}$ and the point $\mathbf{Q}$. Consider further the two 3-vectors:

$$\mathbf{A} = \begin{pmatrix} L_5 \\ L_6 \\ L_4 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} L_2 \\ L_3 \\ L_1 \end{pmatrix}$$

for which the Plücker constraint holds: $\mathbf{B}^\mathsf{T}\mathbf{A} = 0$. $\mathbf{A}$ represents the point at infinity of the line. The Plücker matrix can be written as:

$$[\mathbf{L}]_\times = \begin{pmatrix} 0 & -L_4 & L_6 & -L_2 \\ L_4 & 0 & -L_5 & -L_3 \\ -L_6 & L_5 & 0 & -L_1 \\ L_2 & L_3 & L_1 & 0 \end{pmatrix} = \begin{pmatrix} [\mathbf{A}]_\times & -\mathbf{B} \\ \mathbf{B}^\mathsf{T} & 0 \end{pmatrix}$$

The calibration problem considered in this report is to compute projection rays for all pixels of a camera, from images of planar calibration grids in unknown positions. We assume that dense point correspondences are given, i.e. for (many) pixels, we are able to determine the points on the calibration grids that are seen in that pixel. Computed projection rays will be constrained to cut the camera axis. The coordinate system in which calibration will be expressed, is that of the first calibration grid. Calibration thus consists in computing the position of the camera axis and of the projection rays, in that coordinate system. The proposed approach proceeds by first estimating the camera axis and the pose of all grids but the first one.

3 What can be Done with Two Views of Calibration Grids?

Consider some pixel and let $\mathbf{Q}$ and $\mathbf{Q}'$ be the corresponding points on the two calibration grids, given as 3D points in the grids' local coordinate systems. Since we consider planar grids, we impose $Q_3 = Q'_3 = 0$.

We have the following constraint on the pose of the second grid $(\mathsf{R}', \mathbf{t}')$ as well as the unknown camera axis $\mathbf{L}$: the line spanned by $\mathbf{Q}$ and $\mathbf{Q}'$ cuts $\mathbf{L}$, hence is coplanar with it. Hence, for the correct pose and camera axis, we must have:

$$\mathbf{Q}^\mathsf{T}[\mathbf{L}]_\times \begin{pmatrix} \mathsf{R}' & \mathbf{t}' \\ \mathbf{0}^\mathsf{T} & 1 \end{pmatrix} \mathbf{Q}' = 0$$

Hence:

$$\begin{pmatrix} Q_1 \\ Q_2 \\ Q_4 \end{pmatrix}^\mathsf{T} \begin{pmatrix} 0 & -L_4 & L_6 & -L_2 \\ L_4 & 0 & -L_5 & -L_3 \\ L_2 & L_3 & L_1 & 0 \end{pmatrix} \begin{pmatrix} \bar{\mathsf{R}}' & \mathbf{t}' \\ \mathbf{0}^\mathsf{T} & 1 \end{pmatrix} \begin{pmatrix} Q'_1 \\ Q'_2 \\ Q'_4 \end{pmatrix} = 0$$

where $\bar{\mathsf{R}}'$ refers to the 3×2 submatrix of R' containing only the first and the second rows. We thus have the following 3×3 tensor that can be estimated linearly from point correspondences:

$$\mathsf{F} \sim \begin{pmatrix} 0 & -L_4 & L_6 & -L_2 \\ L_4 & 0 & -L_5 & -L_3 \\ L_2 & L_3 & L_1 & 0 \end{pmatrix} \begin{pmatrix} \bar{\mathsf{R}}' & \mathbf{t}' \\ \mathbf{0}^\mathsf{T} & 1 \end{pmatrix} \tag{1}$$

It has only 7 degrees of freedom (9 - 1 for scale, -1 for rank-deficiency) so the 10 unknowns (4 for the camera axis, 3 for R' and 3 for $\mathbf{t}'$) can not be recovered from it.

We now look at what can actually be recovered from F. Let us first notice that its left null-vector is $(L_3, -L_2, L_4)^\mathsf{T}$ (it truly is the null-vector, as can be easily verified when taking into account the Plücker constraint). We thus can recover 2 of the 4 parameters of the camera axis. That null-vector contains actually the coordinates of the camera axis' intersection with the first grid (in plane coordinates). Its 3D coordinates are given by $(L_3, -L_2, 0, L_4)^\mathsf{T}$. Similarly, the right null-vector of F gives the plane coordinates of the axis' intersection with the second grid. Besides this F also gives constraints on R' and $\mathbf{t}'$. For example R' can be extracted up to 2 to 4 solutions. We will later observe that once we locally shift the intersection points, between the camera axis and calibration grids, to the origins of the respective grids the vector $\mathbf{t}'$ will lie on the camera axis. Inspite of all these additional constraints, arising from axial geometry, two views of calibration grids are not sufficient to uniquely extract R' and $\mathbf{t}'$. Thus we use three calibration grids as described below.

4 Full Calibration using Three Views of Calibration Grids

Let $\mathbf{Q}$, $\mathbf{Q}'$, $\mathbf{Q}''$ refer to the grid points corresponding to a single pixel in the three grids. The poses of the grids are $(\mathsf{I}, \mathbf{0})$, $(\mathsf{R}', \mathbf{t}')$ and $(\mathsf{R}'', \mathbf{t}'')$ respectively. Since the three points $\mathbf{Q}$, $\mathbf{Q}'$ and $\mathbf{Q}''$ are collinear we use this constraint to extract the poses of the calibration grids [15]. Every 3×3 submatrix of the following 4×3 matrix has zero subdeterminant.

$$\begin{pmatrix} \mathbf{Q} & \begin{pmatrix} \mathsf{R}' & \mathbf{t}' \\ \mathbf{0}^\mathsf{T} & 1 \end{pmatrix} \mathbf{Q}' & \begin{pmatrix} \mathsf{R}'' & \mathbf{t}'' \\ \mathbf{0}^\mathsf{T} & 1 \end{pmatrix} \mathbf{Q}'' \end{pmatrix}$$

The submatrices constructed by removing the first and the second rows lead to the constraints $\sum C_i T1_i = 0$ and $\sum C_i T2_i = 0$ respectively (as described in Table 1). These are nothing but homogeneous linear systems of the form $AX = 0$. The unknown vector X is formed from the 14 variables (C_i). Each of these variables are coupled coefficients of the poses of the grids. The matrix A is constructed by stacking the trilinear tensors $T1$ and $T2$, which can be computed from the coordinates of $\mathbf{Q}$, $\mathbf{Q}'$ and $\mathbf{Q}''$. In future when we refer to the rank of a linear system $AX = 0$, we refer to the rank of the matrix A. The rank has to be one less than the number of variables to estimate them uniquely upto a scale. For example, each of the above linear systems must have a rank of 13 to estimate the coefficients (C_i) uniquely. These systems were used to calibrate completely non-central cameras [12]. However in the case of axial cameras, these systems were found to have a rank of 12. This implies that the solution can not be obtained uniquely. In order to resolve this ambiguity we will need more constraints.

4.1 Intersection of axis and calibration grids

Using the technique described earlier we compute the intersection of the camera axis with the three grids at $\mathbf{a}$,$\mathbf{b}$ and $\mathbf{c}$ respectively. We translate the local grid coordinates such that these

i	Motion (C_i)	$T1_i$	$T2_i$	i	Motion (C_i)	$T1_i$	$T2_i$
1	R'_{31}	$Q_2Q'_1Q''_4$	$Q_1Q'_1Q''_4$	13	$R'_{22}R''_{32} - R'_{32}R''_{22}$	$Q_4Q'_2Q''_2$	0
2	R'_{32}	$Q_2Q'_2Q''_4$	$Q_1Q'_2Q''_4$	14	$R'_{11}t''_3 - R'_{31}t''_1$	0	$Q_4Q'_1Q''_4$
3	R''_{31}	$-Q_2Q'_4Q''_1$	$-Q_1Q'_4Q''_1$	15	$R'_{12}t''_3 - R'_{32}t''_1$	0	$Q_4Q'_2Q''_4$
4	R''_{32}	$-Q_2Q'_4Q''_2$	$-Q_1Q'_4Q''_2$	16	$R'_{21}t''_3 - R'_{31}t''_2$	$Q_4Q'_1Q''_4$	0
5	$t'_3 - t''_3$	$Q_2Q'_4Q''_4$	$Q_1Q'_4Q''_4$	17	$R'_{22}t''_3 - R'_{32}t''_2$	$Q_4Q'_2Q''_4$	0
6	$R'_{11}R''_{31} - R'_{31}R''_{11}$	0	$Q_4Q'_1Q''_1$	18	$R''_{11}t'_3 - R''_{31}t'_1$	0	$-Q_4Q'_4Q''_1$
7	$R'_{11}R''_{32} - R'_{31}R''_{12}$	0	$Q_4Q'_1Q''_2$	19	$R''_{12}t'_3 - R''_{32}t'_1$	0	$-Q_4Q'_4Q''_2$
8	$R'_{12}R''_{31} - R'_{32}R''_{11}$	0	$Q_4Q'_2Q''_1$	20	$R''_{21}t'_3 - R''_{31}t'_2$	$-Q_4Q'_4Q''_1$	0
9	$R'_{12}R''_{32} - R'_{32}R''_{12}$	0	$Q_4Q'_2Q''_2$	21	$R''_{22}t'_3 - R''_{32}t'_2$	$-Q_4Q'_4Q''_2$	0
10	$R'_{21}R''_{31} - R'_{31}R''_{21}$	$Q_4Q'_1Q''_1$	0	22	$t'_1t''_3 - t'_3t''_1$	0	$Q_4Q'_4Q''_4$
11	$R'_{21}R''_{32} - R'_{31}R''_{22}$	$Q_4Q'_1Q''_2$	0	23	$t'_2t''_3 - t'_3t''_2$	$Q_4Q'_4Q''_4$	0
12	$R'_{22}R''_{31} - R'_{32}R''_{21}$	$Q_4Q'_2Q''_1$	0				

Table 1: Trifocal tensor in the generic calibration of completely non-central cameras.

intersection points become their respective origins. Without loss of generality we continue to use the same notations after the transformations.

$$\mathbf{Q} \longleftarrow \mathbf{Q} - \mathbf{a},\ \mathbf{Q}' \longleftarrow \mathbf{Q}' - \mathbf{b},\ \mathbf{Q}'' \longleftarrow \mathbf{Q}'' - \mathbf{c},$$

We can obtain a collinearity constraint by putting these origins in the same coordinate system. Every 3×3 subdeterminant of the following 4×3 matrix vanishes.

$$\begin{pmatrix} \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} & \begin{pmatrix} \mathsf{R}' & \mathbf{t}' \\ \mathbf{0}^\mathsf{T} & 1 \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} & \begin{pmatrix} \mathsf{R}'' & \mathbf{t}'' \\ \mathbf{0}^\mathsf{T} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} 0 & t'_1 & t''_1 \\ 0 & t'_2 & t''_2 \\ 0 & t'_3 & t''_3 \\ 1 & 1 & 1 \end{pmatrix}$$

The camera axis passes through O, $\mathbf{t}'$ and $\mathbf{t}''$. This enables us to express $\mathbf{t}''$ as a multiple of $\mathbf{t}'$ using some scalar Δ: $\mathbf{t}'' = \Delta\mathbf{t}'$. As a result, the variables C_{22} and C_{23} from Table 1 disappear.

$$C_{22} = t'_1t''_3 - t'_3t''_1 = t'_1\Delta t'_3 - t'_3\Delta t'_1 = 0$$

$$C_{23} = t'_2t''_3 - t'_3t''_2 = t'_2\Delta t'_3 - t'_3\Delta t'_2 = 0$$

On disappearing, C_{22} and C_{23} reduce the size of the linear systems $\sum C_iT1_i = 0$ and $\sum C_iT2_i = 0$ each by one. Inspite of this reduction there still exists a rank deficiency of 2 in both these systems. The rank of each of these systems is 11 with 13 nonzero coefficients to be estimated. In the next section we provide the details of the usage of a coplanarity constraint, which exists in axial cameras, to remove the degeneracy problems.

4.2 Coplanarity constraints in axial cameras

The camera axis cuts all the projection rays. As observed earlier both $\mathbf{O}$ and $\mathbf{t}'$ lie on the camera axis. Along with these two points, we consider two grid points $\mathbf{Q}'$ and $\mathbf{Q}''$ lying on a single projection ray. Since these four points are coplanar, the determinant of the following

4×4 matrix disappears.

$$\left(\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad \begin{pmatrix} t'_1 \\ t'_2 \\ t'_3 \\ 1 \end{pmatrix} \quad \begin{pmatrix} \mathsf{R}' & \mathbf{t}' \\ \mathbf{0}^\mathsf{T} & 1 \end{pmatrix} \mathbf{Q}' \quad \begin{pmatrix} \mathsf{R}'' & \mathbf{\Delta} t' \\ \mathbf{0}^\mathsf{T} & 1 \end{pmatrix} \mathbf{Q}'' \right)$$

The corresponding constraint is a linear system $\sum \alpha_{ij} Q'_i Q''_j = 0$ (see table 2). Note that Q'_4 and Q''_4 are not present because of the three zeros in the first column. We can solve this linear system to computer the solutions for α_{ij}. We expand the above linear system and do some algebraic manipulation.

$$\alpha_{11} Q'_1 Q''_1 + \alpha_{12} Q'_1 Q''_2 + \alpha_{21} Q'_2 Q''_1 + \alpha_{22} Q'_2 Q''_2 = 0$$

$$Q_4(\alpha_{11} Q'_1 Q''_1 + \alpha_{12} Q'_1 Q''_2 + \alpha_{21} Q'_2 Q''_1 + \alpha_{22} Q'_2 Q''_2) = 0$$

$$Q_4 Q'_2 Q''_2 = -\frac{\alpha_{11}}{\alpha_{22}} Q_4 Q'_1 Q''_1 - \frac{\alpha_{12}}{\alpha_{22}} Q_4 Q'_1 Q''_2 - \frac{\alpha_{21}}{\alpha_{22}} Q_4 Q'_2 Q''_1$$

This will enable us to represent both $T2_9$ and $T1_{13}$, from the earlier systems, in terms of other variables in the tensors $T1$ and $T2$ respectively.

$$T2_9 = -\frac{\alpha_{11}}{\alpha_{22}} T2_6 - \frac{\alpha_{12}}{\alpha_{22}} T2_7 - \frac{\alpha_{21}}{\alpha_{22}} T2_8$$

$$T1_{13} = -\frac{\alpha_{11}}{\alpha_{22}} T1_{10} - \frac{\alpha_{12}}{\alpha_{22}} T1_{11} - \frac{\alpha_{21}}{\alpha_{22}} T1_{12}$$

Using the above relation we obtain two new constraints given by $\sum A_i A1_i = 0$ and $\sum A_i A2_i = 0$. Note that each of these constraints are linear systems with 12 nonzero coefficients each. Both of them have a rank of 11 and thereby producing unique solutions for their coefficients (A_i). The individual elements in the poses of the grids are extracted from these coupled coefficients using orthonormality constraints of the rotation matrix [15].

i	j	α_{ij}
1	1	$t'_1(R'_{2,1}R''_{3,1} - R''_{2,1}R'_{3,1}) - t'_2(R'_{1,1}R''_{3,1} - R''_{1,1}R'_{3,1}) + t'_3(R'_{1,1}R''_{2,1} - R''_{1,1}R'_{2,1})$
1	2	$t'_1(R'_{2,1}R''_{3,2} - R''_{2,2}R'_{3,1}) - t'_2(R'_{1,1}R''_{3,2} - R''_{1,2}R'_{3,1}) + t'_3(R'_{1,1}R''_{2,2} - R''_{1,2}R'_{2,1})$
2	1	$t'_1(R'_{2,2}R''_{3,1} - R''_{2,1}R'_{3,2}) - t'_2(R'_{1,2}R''_{3,1} - R''_{1,1}R'_{3,2}) + t'_3(R'_{1,2}R''_{2,1} - R''_{1,1}R'_{2,2})$
2	2	$t'_1(R'_{2,2}R''_{3,2} - R''_{2,2}R'_{3,2}) - t'_2(R'_{1,2}R''_{3,2} - R''_{1,2}R'_{3,2}) + t'_3(R'_{1,2}R''_{2,2} - R''_{1,2}R'_{2,2})$

Table 2: Bifocal tensor from the coplanarity constraint on $\mathbf{O}$, $\mathbf{t}'$, $\mathbf{Q}'$ and $\mathbf{Q}''$.

5 Alternative approach for Axial cameras using subspace techniques.

The above technique involves solving several linear systems. In the case of real cameras, due to the presence of noise, the above technique was not stable. In the following we propose a different approach using subspaces. Table 4 gives the details for a 3D scenario with the planar calibration boards.

i	Motion (A_i)	$A1_i$	$A2_i$	i	Motion (A_i)	$A1_i$	$A2_i$
1	R'_{31}	$Q_2Q'_1Q''_4$	$Q_1Q'_1Q''_4$	11	$C_{12} - \frac{\alpha_{21}}{\alpha_{22}}C_{13}$	$Q_4Q'_2Q''_1$	0
2	R'_{32}	$Q_2Q'_2Q''_4$	$Q_1Q'_2Q''_4$	12	$\Delta(R'_{11}t'_3 - R'_{31}t'_1)$	0	$Q_4Q'_1Q''_4$
3	R''_{31}	$-Q_2Q'_4Q''_1$	$-Q_1Q'_4Q''_1$	13	$\Delta(R'_{12}t'_3 - R'_{32}t'_1)$	0	$Q_4Q'_2Q''_4$
4	R''_{32}	$-Q_2Q'_4Q''_2$	$-Q_1Q'_4Q''_2$	14	$\Delta(R'_{21}t'_3 - R'_{31}t'_2)$	$Q_4Q'_1Q''_4$	0
5	$t'_3 - t''_3$	$Q_2Q'_4Q''_4$	$Q_1Q'_4Q''_4$	15	$\Delta(R'_{22}t'_3 - R'_{32}t'_2)$	$Q_4Q'_2Q''_4$	0
6	$C_6 - \frac{\alpha_{11}}{\alpha_{22}}C_9$	0	$Q_4Q'_1Q''_1$	16	$R''_{11}t'_3 - R''_{31}t'_1$	0	$-Q_4Q'_4Q''_1$
7	$C_7 - \frac{\alpha_{12}}{\alpha_{22}}C_9$	0	$Q_4Q'_1Q''_2$	17	$R''_{12}t'_3 - R''_{32}t'_1$	0	$-Q_4Q'_4Q''_2$
8	$C_8 - \frac{\alpha_{21}}{\alpha_{22}}C_9$	0	$Q_4Q'_2Q''_1$	18	$R''_{21}t'_3 - R''_{31}t'_2$	$-Q_4Q'_4Q''_1$	0
9	$C_{10} - \frac{\alpha_{11}}{\alpha_{22}}C_{13}$	$Q_4Q'_1Q''_1$	0	19	$R''_{22}t'_3 - R''_{32}t'_2$	$-Q_4Q'_4Q''_2$	0
10	$C_{11} - \frac{\alpha_{12}}{\alpha_{22}}C_{13}$	$Q_4Q'_1Q''_2$	0				

Table 3: Trifocal tensor for the generic calibration of axial cameras.

i	C_i	V_i	W_i
1	$Q_1Q'_1Q''_4$	0	$R'_{3,1}$
2	$Q_1Q'_2Q''_4$	0	$R'_{3,2}$
3	$Q_1Q'_4Q''_1$	0	$-R''_{3,1}$
4	$Q_1Q'_4Q''_2$	0	$-R''_{3,2}$
5	$Q_1Q'_4Q''_4$	0	$t'_3 - t''_3$
6	$Q_2Q'_1Q''_4$	$R'_{3,1}$	0
7	$Q_2Q'_2Q''_4$	$R'_{3,2}$	0
8	$Q_2Q'_4Q''_1$	$-R''_{3,1}$	0
9	$Q_2Q'_4Q''_2$	$-R''_{3,2}$	0
10	$Q_2Q'_4Q''_4$	$t'_3 - t''_3$	0
11	$Q_4Q'_1Q''_1$	$-R''_{2,1}R'_{3,1} + R'_{2,1}R''_{3,1}$	$R'_{1,1}R''_{3,1} - R''_{1,1}R'_{3,1}$
12	$Q_4Q'_1Q''_2$	$R'_{2,1}R''_{3,2} - R''_{2,2}R'_{3,1}$	$R'_{1,1}R''_{3,2} - R''_{1,2}R'_{3,1}$
13	$Q_4Q'_1Q''_4$	$R'_{2,1}t''_3 - t''_2R'_{3,1}$	$R'_{1,1}t''_3 - t''_1R'_{3,1}$
14	$Q_4Q'_2Q''_1$	$-R''_{2,1}R'_{3,2} - R''_{2,2}R'_{3,2}$	$R'_{1,2}R''_{3,1} - R''_{1,1}R'_{3,2}$
15	$Q_4Q'_2Q''_2$	$R'_{2,2}R''_{3,2} - R''_{2,2}R'_{3,2}$	$R'_{1,2}R''_{3,2} - R''_{1,2}R'_{3,2}$
16	$Q_4Q'_2Q''_4$	$R'_{2,2}t''_3 - t''_2R'_{3,2}$	$R'_{1,2}t''_3 - t''_1R'_{3,2}$
17	$Q_4Q'_4Q''_1$	$-R''_{2,1}t'_3 + t'_2R''_{3,1}$	$t'_1R''_{3,1} - R''_{1,1}t'_{3,1}$
18	$Q_4Q'_4Q''_2$	$t'_2R''_{3,2} - R''_{2,2}t'_3$	$t'_1R''_{3,2} - R''_{1,2}t'_{3,1}$
19	$Q_4Q'_4Q''_4$	$t'_2t''_3 - t''_2t'_3$	$t'_1t''_3 - t''_1t'_3$

Table 4: Coupled variables in the 3D general camera with a planar board pattern. This table is taken from [14].

On applying the constraints of generic calibration, which is nothing but the collinearity constraint on the points lying on the projection ray, we obtain equations of the form $\sum C_i V_i = 0$ and $\sum C_i W_i = 0$. There are two of the four possible constraints that can be obtained from the collinearity condition. These constraints are homogeneous linear systems of the form $AX = 0$. Our goal is to first compute the vector X, which can be V or W in this case, and then extract the individual motion variables. If the rank of A, which has a dimension of $m \times n$ $(m > n)$, is $n - 1$ then there exists a unique solution to X up to a scale. The required solution is equal to the eigenvector corresponding to the least eigenvalue of $A^T A$. Using simulations we found that the above system always has a rank deficiency of two instead of one, when we apply it to an axial camera. Consequently we can only obtain the required solution in a subspace spanned by two vectors. In other words our required solution is present in the subspace spanned by the two eigenvectors corresponding to the last two eigenvalues of $A^T A$. Let V and V^P represent the two eigenvectors for the first constraint. Similarly we use W^P in addition to W for the second constraint. We represent our required solution in terms of unknown parameters λ_v and λ_w. Thus by computing these parameters we will be able to remove the degeneracy problem. In the following we describe a method to compute these two parameters. Now the collinearity constraints are shown below.

$$\sum_{i=1}^{i=14} C_i(V_i + \lambda_v V_i^P) = 0$$

$$\sum_{i=1}^{i=14} C_i(W_i + \lambda_w W_i^P) = 0$$

From the table we can observe that the first five nonzero elements represented by $V + \lambda_v V_P$ and $W + \lambda_w W^P$ are the same. Using this relation we can compute λ_v and λ_w if the rank of the following matrix is 3.

$$\begin{bmatrix} V_6 & V_6^P & W_1 & W_1^P \\ V_7 & V_7^P & W_2 & W_2^P \\ V_8 & V_8^P & W_3 & W_3^P \\ V_9 & V_9^P & W_4 & W_4^P \\ V_{10} & V_{10}^P & W_5 & W_5^P \end{bmatrix}$$

However the rank of the above matrix was always one.

$$\begin{bmatrix} V_6 \\ V_7 \\ V_8 \\ V_9 \\ V_{10} \end{bmatrix} + \lambda_v \begin{bmatrix} V_6^P \\ V_7^P \\ V_8^P \\ V_9^P \\ V_{10}^P \end{bmatrix} \sim \begin{bmatrix} V_6 \\ V_7 \\ V_8 \\ V_9 \\ V_{10} \end{bmatrix}$$

$$\begin{bmatrix} W_6 \\ W_7 \\ W_8 \\ W_9 \\ W_{10} \end{bmatrix} + \lambda_w \begin{bmatrix} W_6^P \\ W_7^P \\ W_8^P \\ W_9^P \\ W_{10}^P \end{bmatrix} \sim \begin{bmatrix} W_6 \\ W_7 \\ W_8 \\ W_9 \\ W_{10} \end{bmatrix}$$

Using the above relation we compute the variables up to the common scale and represent it by $V' + \lambda_v VP'$ and $W' + \lambda_w WP'$.

$$R'_{3,1} \sim (V'_6 + \lambda_v V'^P_6) \sim V'_6$$

$$R'_{3,2} \sim (V'_7 + \lambda_v {V'}^P_7) \sim V'_7$$
$$R''_{3,1} \sim (-V'_8 - \lambda_w {V'}^P_8) \sim -V'_8$$
$$R''_{3,2} \sim (-V'_9 - \lambda_w {V'}^P_9) \sim -V'_9$$
$$\lambda R'_{3,1} = V'_6$$
$$\lambda R'_{3,2} = V'_7$$
$$\lambda R''_{3,1} = -V'_8$$
$$\lambda R''_{3,2} = -V'_9$$

Next we construct a system of linear equations for evaluating $R'_{2,1}$, $R'_{2,2}$,$R''_{2,1}$,$R''_{2,2}$ and λ_v.

$$\begin{bmatrix} -V'_8 & 0 & -V'_6 & 0 & -{W'}^P_{11} \\ -V'_9 & 0 & 0 & -V'_6 & -{W'}^P_{12} \\ 0 & -V'_8 & -V'_7 & 0 & -{W'}^P_{14} \\ 0 & -V'_9 & 0 & -V'_7 & -{W'}^P_{15} \end{bmatrix} \begin{bmatrix} R'_{2,1} \\ R'_{2,2} \\ R''_{2,1} \\ R''_{2,2} \\ \lambda_v \end{bmatrix} = \begin{bmatrix} V'_{11} \\ V'_{12} \\ V'_{14} \\ V'_{15} \end{bmatrix}$$

The rank of the above system is 4. So the solution is obtained in a subspace spanned by two vectors a and b.

$$\begin{bmatrix} R'_{2,1} \\ R'_{2,2} \\ R''_{2,1} \\ R''_{2,2} \\ \lambda_v \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} + l_1 \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ a_6 \end{bmatrix}$$

Then use the orthonormal properties of the rotation matrix to compute the linear factor l_1, which is then used to compute λ_v. Similarly we can compute λ_w. Once we compute λ_v and λ_w we can follow the earlier algorithm given in section 5.2 of [14] to extract the individual motion parameters. Note that this algorithm is also applicable for completely general cameras, or in other words the noncentral cameras. In such cases we found that the λ_v and λ_w were zeros.

6 Bundle Adjustment Formulation

We give the details of a bundle adjustment which refines the estimated camera axis and poses of the calibration grids. This is similar to our earlier method [12], except that we have an additional constraint coming from the camera axis. The bundle adjustment is done by minimizing the distance between the grid points and the corresponding projection rays. The cost function is given below.

$$Cost = \sum_{i=1}^{n} \sum_{j=1} (\mathbf{A} + \lambda_i \mathbf{D} + \mu_{ji} \mathbf{D}_i - [\mathsf{R}_j \mathbf{T}_j] \mathbf{Q}_{ji})$$

- $(\mathbf{A}, \mathbf{D})$ - represents the axis (point, direction)
- $\mathbf{D}_i$ - unit direction vector of the i_{th} projection ray
- λ_i - parameter selecting the intersection of the i_{th} ray and the axis

- $\mathbf{Q}_{ji}$ - grid point on the j_{th} grid lying the i_{th} ray
- μ_{ji} - parameter selecting the point on the i_{th} ray closest to Q_j
- $(\mathsf{R}_j, \mathbf{T}_j)$ - pose of the calibration grid

7 Axial Catadioptric Configurations

Our formulation can classify a given camera into either axial or not. For example on applying our method on axial data we obtain unique solutions. On the other hand, a completely non-central camera will lead to an inconsistent (no solution), whereas a central camera will produce a rank deficient system (ambiguous solutions). Thus our technique produces unique solutions only for axial configurations. This can be used as a simple test in simulations to study the nature of complex catadioptric arrangements (as shown in Figure 2(a)). Since axial cameras are less restrictive than central cameras, they can be easily constructed using various combinations of mirrors and lenses. For example there are very few central configurations [2] (also see Table 5). Furthermore these configurations are difficult to build and maintain. For example, in a central catadioptric camera with hyperbolic mirror and perspective camera, the optical center has to be placed precisely on one of the mirror's focal points. On the other hand, the optical center can be anywhere on the mirror axis to have an axial geometry.

mirror	ctrl (pers)	axial (pers)	nctrl (pers)	ctral (ortho)	axial (ortho)	nctrl (ortho)
hyperbolic	o=f	o $\in$ MA	o $\notin$ MA	-	OA $\parallel$ MA	OA $\nparallel$ MA
spherical	-	always	-	-	always	-
parabolic	-	o $\in$ MA	o $\notin$ MA	OA $\parallel$ MA	-	OA $\nparallel$ MA
elliptic	o = f	o $\in$ MA	o $\notin$ MA	-	OA $\parallel$ MA	OA $\nparallel$ MA
cone	-	o $\in$ MA	o $\notin$ MA	-	OA $\parallel$ MA	OA $\nparallel$ MA
planar	always	-	planar	-	-	-
mir-rot	-	always	-	-	always	-

Table 5: Catadioptric configurations. Notations: ctrl (pers) - central configuration with perspective camera, nctrl (ortho) - non-central configuration with orthographic camera, mir-rot - mirror obtained by rotating a planar curve about the optical axis, o - optical center of the perspective camera, f - focus of the mirror, MA - major axis of mirror, OA - optical axis of the camera, = refers to same location, $\in$-lies on, $\parallel$-parallel, $\nparallel$-not parallel.

8 Experiments

8.1 Simulation

We started with perfect axial configurations for three scenarios (as shown in Figures 2(a), (b) and (c)) and gradually change the configurations to make them non-central. We quantify this change from the perfect axial configuration as disparity. For example, in Figure 2(a), the disparity represents the distance between the optical center of the perspective camera and the orthographic camera axis passing through the center of the sphere. This optical center is

initially at a distance of 3 units from the center of the sphere (which is of radius 1 unit). In Figure 2(b), the disparity represents the distance between the optical center of the perspective camera and the major axis of the hyperboloid. Initially the optical center is at a distance of 5 units from the tip of the hyperboloid, whose two radii are 5 and 10 units. In Figure 2(c), the disparity represents the distance between the optical center of the third camera and the line joining the first two cameras. The distance between two consecutive centers of the cameras is 40 units. We calibrate these systems in the presence of disparities. We compute the mean angular error between the original and the reconstructed projection rays in Figure 2(d). Note that the the mean angular error (given in radians) reaches zero only at the precise axial configuration.

8.2 Stereo camera

We tested our axial calibration algorithm for a stereo system. We captured three images of a calibration grid using a stereo system. The focus is to reconstruct the projection rays for both the cameras in the same generic framework using our axial calibration algorithm. The image of the combined system is formed by concatenating the images from the two cameras. Here the camera axis is the line joining the two optical centers (see Figure 4(a)). The experiments in section 8.1 was performed using the method provided in section 4. However this algorithm is very sensitive to noise as shown in Figure 2(d). In order to deal with this problem we developed a technique using subspaces, which was described in section 5. This technique is more stable to noise and works well for stereo data.

Once we compute the poses of the grids we can compute the rays corresponding to individual cameras in the stereo system. The reconstructed rays are shown in Figure 3(a). We compute the axis using the constraints given in section 4. Since the axis was reconstructed using one sample, which consists of a set of 12 rays producing the least ray-point error in RANSAC, the axis does not pass through all the rays. The Ray-Point RMS error is 0.04 in percent with respect to the overall size of the scene. We compute the intersection points, or the points closest to the projection rays on the axis. Each of these projection rays is forced to pass through these closest points on the axis and the ray directions are recomputed using the 3D points from the calibration grids. We then clustered the rays into two sets, based on the points on the axis, corresponding to the two pinhole cameras. The centroids of these clusters correspond to the two centers of the stereo systems. Again the rays are forced to pass through their corresponding centers and the ray directions are recomputed. We used an axial bundle adjustment algorithm (see section 6). The calibration statistics for the stereo system is given in Table 6. The final Ray-Point RMS error is 0.07 in percent with respect to the overall size of the scene. Note that the ray-point RMS error increases after fitting the axis. This is due to the overfitting of the noise before forcing the axial constraint.

In order to evaluate the calibration, we compared results with those obtained by plane-based calibration [16, 18], that used the knowledge that the two cameras are pinholes. In both, our axial calibration, and plane-based calibration, the first grid was used to fix the global coordinate system. We can thus compare the estimated poses of the other two grids for the two methods. This is done for both, the rotational and translational parts of the pose. As for rotation, we measure the angle (in radians) of the relative rotation between the rotation matrices given by the two methods, see columns R_i in Table 7). As for translation, we measure the distance between the estimated 3D positions of the grids' centers of gravity (columns t_i in Table 7) expressed in percent, relative to the scene size. Here, plane-based calibration is done separately for each camera, leading to the two rows of Table 7.

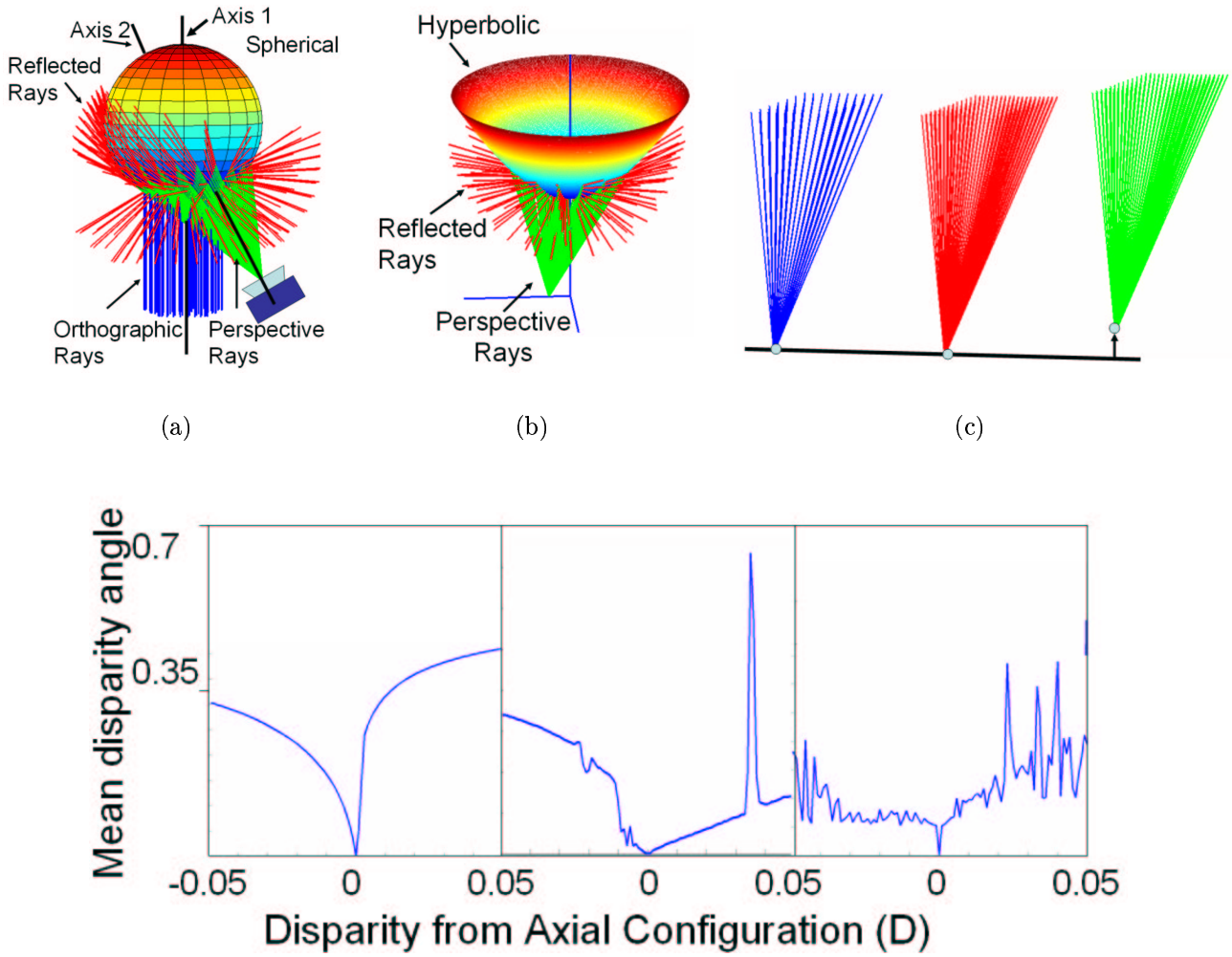
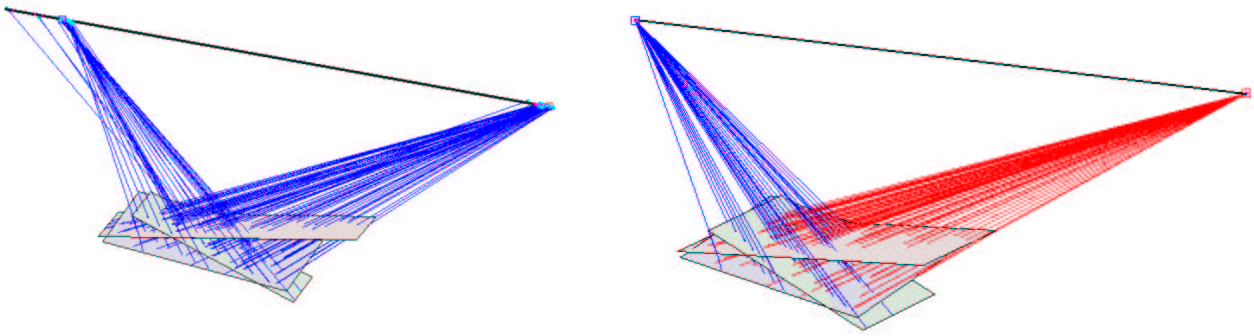


Figure 2: Test for axial configuration. (a) Catadioptric (spherical mirror+pers.camera+ortho.camera): becomes non-central when the two optical centers and the sphere center are not collinear (as shown).(b) Catadioptric (Hyperbolic mirror+pers.camera): becomes non-central if the optical center is not on the axis of the hyperbolic mirror (as shown). (c) Tristereo when one of the cameras is axially misplaced (as shown). (d) shows the mean angular error between the original and reconstructed projection rays w.r.t disparity. The graphs shown in left, middle and right correspond to scenarios in (a), (b) and (c) respectively (see text for more details).

Images	Rays	Points	RMS before fitting the axis	RMS after fitting the axis
3	481	3x481	0.04	0.07

Table 6: Bundle adjustment statistics for different cameras. RMS is the root-mean-square residual error of the bundle adjustment (ray-point distances). It is given in percent, relative to the overall size of the scene (largest pairwise distance between points on calibration grids).



(a) (b)

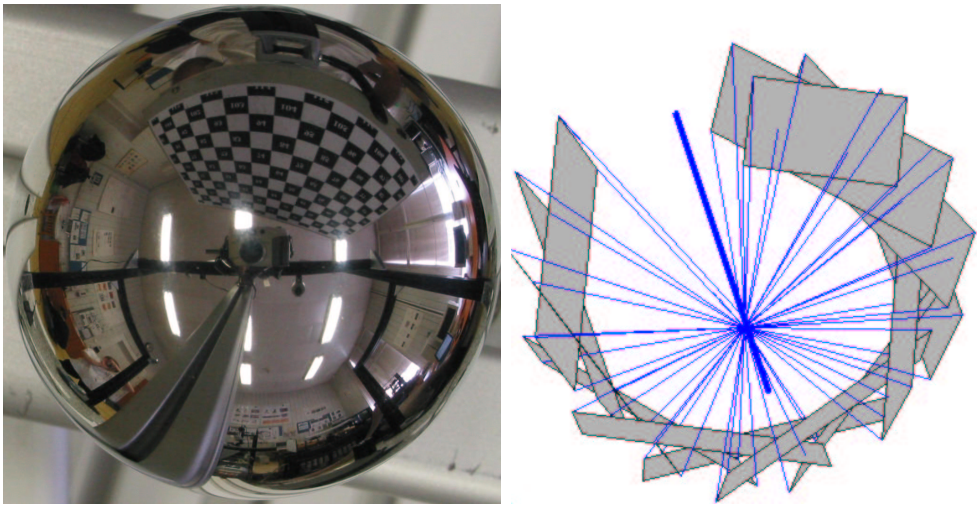
Figure 3: Calibration of a stereo system using axial algorithm. (a) Reconstructed projection rays. Note that the rays corresponding to each of the pinhole cameras do not pass through a single point. This is because the rays are computed using a generic algorithm for axial cameras. (b) Clustered rays, corresponding to a stereo system.

Camera	R_2	R_3	t_2	t_3	Center
1	0.0438	0.0342	0.46	0.75	3.35
2	0.0603	0.0491	0.37	0.61	1.96

Table 7: Evaluation of axial stereo calibration relative to plane based calibration. The comparison is made on the pose estimation of the calibration grids and the estimation of the camera centers. See text for more details.

8.3 Spherical catadioptric cameras

We calibrated a spherical catadioptric camera and extracted the camera axis. First we use a central generic calibration for three calibration grids. Once the poses of the grids are estimated we compute the axis using the constraints, provided in section 4. This enables us to obtain an initial estimate for the axis and the projection rays. Now we force the projection rays to pass through the camera axis. Using this partial calibration, we use pose estimation to incrementally compute the pose of newer grids. The calibration grid captured by a spherical catadioptric camera is shown in Figure 4(b). We estimated the pose of several grids on a turntable sequence using the calibration. The grid positions and the axis are shown in Figure 4(c). Finally we use an axial bundle adjustment (see section 6). The spherical catadioptric model is very close to a central model than an axial one. All the points, obtained from the intersection of the projection rays and the camera axis, are distributed in a very small region on the camera axis. This can can be approximated as the center of the camera. As a result of this our axial algorithm is not very robust to spherical catadioptric cameras. This is the reason for using central approximation for initializing the calibration algorithm. However in other axial catadioptric cameras, where the intersection points are widely distributed, the axial algorithm will be more robust.



(a) (b)

Figure 4: Calibration of a spherical catadioptric camera. (a) Image of a calibration grid. (b) Estimated poses of several grids along with the camera axis.

9 Conclusions

We studied the theory and proposed a linear calibration algorithm for an intermediate class of cameras called axial cameras. Further line of investigation needs to be carried out to test the accuracy of this approach with respect to parametric and completely non-central approaches.

Acknowledgments: We thank Tomáš Pajdla, Branislav Mičušík and Diana Mateus for the data.

References

[1] D. Aliaga. Accurate Catadioptric Calibration for Real-size Pose Estimation of Room-size Environments, *ICCV*, 2001.

[2] S. Baker and S. Nayar. A theory of catadioptric image formation. *ICCV*, 1998.

[3] H. Bakstein and T. Pajdla. An overview of non-central cameras. *Computer Vision Winter Workshop*, Ljubljana, Slovenia, 2001.

[4] Doron Feldman, Tomas Pajdla and Daphna Weinshall. On the Epipolar Geometry of the Crossed-Slits Projection. *ICCV*, 2003.

[5] M.D. Grossberg and S.K. Nayar. A general imaging model and a method for finding its parameters. *ICCV*, 2001.

[6] R. Gupta and R.I. Hartley. Linear Pushbroom Cameras. *PAMI* 1997.

[7] B. Micusik and T. Pajdla. Autocalibration and 3D Reconstruction with Non-central Catadioptric Cameras. *CVPR*, 2004.

[8] J. Neumann, C. Fermüller, and Y. Aloimonos. Polydioptric Camera Design and 3D Motion Estimation. *CVPR*, 2003.

[9] T. Pajdla. Stereo with oblique cameras. *IJCV*, 2002.

[10] S. Peleg, M. Ben-Ezra, and Y. Pritch. Omnistereo: Panoramic Stereo Imaging. *PAMI*, 2001.

[11] R. Pless. Using Many Cameras as One. In *CVPR*, 2003.

[12] S. Ramalingam, P. Sturm and S.K. Lodha. Towards Complete Generic Camera Calibration. *CVPR*, 2005.

[13] S. Seitz and J. Kim. The Space of All Stereo Images. *IJCV*, 2002.

[14] P. Sturm and S. Ramalingam. A generic calibration concept-theory and algorithms. Research Report 5058, INRIA, 2003.

[15] P. Sturm and S. Ramalingam. A generic concept for camera calibration. *ECCV*, 2004.

[16] P. Sturm and S. Maybank. On Plane-Based Camera Calibration: A General Algorithm, Singularities, Applications. *CVPR*, 1999.

[17] R. Swaminathan, M.D. Grossberg, and S.K. Nayar. A perspective on distortions. *CVPR*, 2003.

[18] Z. Zhang. A flexible new technique for camera calibration. *PAMI*, 2000.

Unité de recherche INRIA Rhône-Alpes
655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)
Unité de recherche INRIA Sophia Antipolis : 2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399