

# Towards Complete Generic Camera Calibration

Srikumar Ramalingam, Peter Sturm, Suresh K. Lodha

▶ **To cite this version:**

Srikumar Ramalingam, Peter Sturm, Suresh K. Lodha. Towards Complete Generic Camera Calibration. IEEE Conference on Computer Vision and Pattern Recognition (CVPR '05), Jun 2005, San Diego, United States. pp.1093-1098, 10.1109/CVPR.2005.347 . inria-00524407

**HAL Id: inria-00524407**

**https://inria.hal.science/inria-00524407v1**

Submitted on 25 May 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# Towards Complete Generic Camera Calibration

Srikumar Ramalingam[1], Peter Sturm[2], and Suresh K. Lodha[1]
[1] Dept. of Computer Science, University of California, Santa Cruz, CA 95064, USA
[2] INRIA Rhône-Alpes, GRAVIR-CNRS, 38330 Montbonnot, France

## Abstract

*We consider the problem of calibrating a highly generic imaging model, that consists of a non-parametric association of a projection ray in 3D to every pixel in an image. Previous calibration approaches for this model do not seem to be directly applicable for cameras with large fields of view and non-central cameras. In this paper, we describe a complete calibration approach that should in principle be able to handle any camera that can be described by the generic imaging model. Initial calibration is performed using multiple images of overlapping calibration grids simultaneously. This is then improved using pose estimation and bundle adjustment-type algorithms. The approach has been applied on a wide variety of central and non-central cameras including fisheye lens, catadioptric cameras with spherical and hyperbolic mirrors, and multi-camera setups. We also consider the question if non-central models are more appropriate for certain cameras than central models.*

## 1. Introduction

This paper is about camera calibration. We adopt a general non-parametric imaging model that consists in associating one projection ray to each individual pixel. By projection ray we refer to the 3D (half-) line along which light travels that falls onto the pixel (here, we neglect point spread and the finite spatial extent of a pixel). Rays may be unconstrained, i.e. they may not intersect in a single point, in which case the camera is called *non-central*. This general model has been used in various works [7, 12, 14, 15, 16, 19, 20, 22, 23, 25, 26], and is best described in [7], where properties other than geometric ones are also considered.

By adopting this model, one may formulate "black-box calibration" and provide algorithms that allow to calibrate any camera (see figure 1 for examples), be it of pinhole type (with or without optical distortions), catadioptric [2, 10], pushbroom [8], or some other acquisition system [15, 20]. Such calibration algorithms have been proposed in [3, 6, 7, 22]. In this paper, we adopt the approach of [22] which allows to perform calibration from three images of a calibration grid, without having to know the motion between the images. To calibrate the complete image with only three images, one would need a calibration grid of appropriate dimensions and shape; especially for omnidirectional cameras (fisheye, catadioptric, etc), this will be cumbersome to produce and handle. In this paper, we propose a similar method, that uses multiple images to accurately and completely calibrate large fields of view.

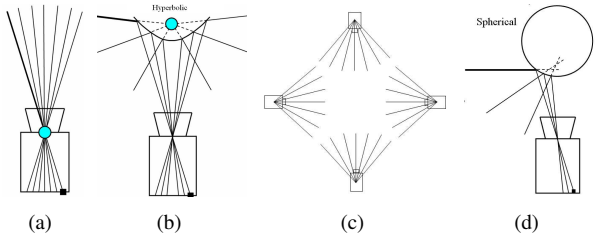



Figure 1. Examples of generic imaging model. (a) pinhole camera, (b) catadioptric with hyperbolic mirror (central), (c) multi-camera, (d) catadioptric with spherical mirror (non-central).

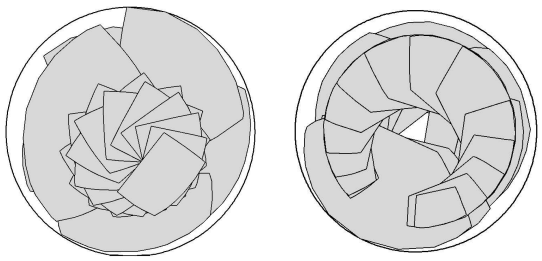


Figure 2. Examples of complete calibration. Left: 23 calibration grids, used in calibrating a fisheye. Right: 24 calibration grids used in calibrating a spherical catadioptric camera.

Our approach works as follows. An initial calibration is done with images of calibration grids that present sufficient overlap. We then recursively incorporate additional images: at each step, we select the image that has the largest overlap with the already calibrated image region. We show how to compute the pose of the associated calibration grid. Then, given the pose, one may compute projection rays for previously uncalibrated pixels, thus enlarging the calibrated image region. This is iterated until all images have been used. We also propose a bundle adjustment method that can be used at any stage of the procedure. This approach and the underlying methods are developed for both, non-central and central models, although the central case is described in more detail here. Besides developing algorithms, we are also interested in the question if for certain cameras it is worth going to a full non-central model, cf. also [1, 11].

This paper is organized as follows. The calibration approach is described in §2 and some variants are proposed in §3. Practical issues are discussed in §4. Experimental results are presented in §5, followed by conclusions in §6.

## 2. Complete Calibration

We first provide an overview of complete generic camera calibration. We take several images of a calibration grid such as to cover the entire image region. Then, matching between image pixels and points on the calibration grids is performed. From such matches, we then compute the pose of each of these grids in a common coordinate system. After this pose computation, a 3D projection ray is computed for each pixel, as follows. For all grid points matching a given pixel, we compute their 3D coordinates (via the pose of the grids). The pixel's projection ray is then simply computed by fitting a straight line to the associated grid points.

For a non-central camera, atleast two grid points per pixel are of course required. If the camera is (assumed to be) central however, a single grid point is enough: as will be seen later, the above stage of pose computation also comprises the estimation of the camera's optical center (in the same coordinate frame as the grids' pose). Thus, we compute projection rays by fitting lines to 3D points, but which are constrained to contain the optical center.

In the following, we describe different parts of our approach in more detail. In this section, we describe the case of central cameras. For conciseness, the non-central case is described more briefly in §3.1. First, we show how to use the images of multiple grids simultaneously, to compute grid pose and the optical center. It is then shown how to compute the pose of additional grids. Refinement of calibration after each step, through bundle adjustment, is then discussed in §2.3.

### 2.1 Calibration using Multiple Grids

Our goal is to obtain the poses of multiple calibration grids w.r.t. a common coordinate system. Let $B_i$ denote the image region covered by the $i_{th}$ calibration grid, for $i = 1 \cdots n$. Let $\cup$ and $\cap$ refer to union and intersection operations respectively. The calibration algorithm is applied to a partial region given by $\cup_{i=2}^{n}(B_1 \cap B_i)$. Once the poses are computed the calibration is extended to a larger region given by $\cup_{i=1}^{n} B_i$.

We now outline the theory behind calibration using multiple grids. Consider one pixel and its associated grid points, with homogeneous coordinates $Q^i = (Q_1^i, Q_2^i, Q_3^i, Q_4^i)^T$, for grids $i = 1 \cdots n$. In the following, we consider planar calibration grids, and thus suppose that $Q_3^i = 0$. Let the unknown grid poses be represented by rotation matrices $R^i$ and translation vectors $t^i$, such that the point $Q^i$, given in local grid coordinates, is mapped to global coordinates via

$$\begin{pmatrix} R^i & t^i \\ 0^T & 1 \end{pmatrix} \begin{pmatrix} Q_1^i \\ Q_2^i \\ 0 \\ Q_4^i \end{pmatrix} \tag{1}$$

Furthermore, let $O = (O_1, O_2, O_3, 1)$ be the coordinates of the camera's optical center. As global coordinate system, we adopt, without loss of generality, the reference frame of the first grid, i.e. $R^1 = I$ and $t^1 = 0$.

We now show how to estimate the unknown grid poses and the optical center. This is based on the following collinearity constraint: with the correct poses, the grid points associated with one pixel, after mapping into the global coordinate system via (1), must be collinear, and in addition, collinear with the optical center. This is because all these points must lie on the pixel's projection ray, i.e. a straight line. Algebraically, this collinearity constraint can be formulated as follows. Consider the matrix containing the coordinates of the collinear points:

$$\begin{pmatrix} O_1 & Q_1^1 & R_{11}^2 Q_1^2 + R_{12}^2 Q_2^2 + t_1^2 Q_4^2 & \cdots \\ O_2 & Q_2^1 & R_{21}^2 Q_1^2 + R_{22}^2 Q_2^2 + t_2^2 Q_4^2 & \cdots \\ O_3 & 0 & R_{31}^2 Q_1^2 + R_{32}^2 Q_2^2 + t_3^2 Q_4^2 & \cdots \\ 1 & Q_4^1 & Q_4^2 & \cdots \end{pmatrix} \tag{2}$$

The collinearity of these points implies that this $4 \times (n+1)$ matrix must be of rank smaller than 3. Consequently, the determinants of all its $3 \times 3$ submatrices must vanish. This gives equations linking calibration point coordinates and the unknowns (camera poses and optical center). On using the first column (optical center) and two other columns with $Q^j$ and $Q^k$ to form a submatrix, we get bilinear equations in terms of calibration point coordinates $Q^j$ and $Q^k$. Hence, we may write the equations in the form:

$$\left(Q^j\right)^T T_{3\times 3} Q^k = 0 \tag{3}$$

This matrix $T$ (a bifocal matching tensor), depends on camera pose and optical center, in a way specific to which $3 \times 3$ submatrix of (2) is considered. Using (3), we estimate such tensors $T$ from available correspondences. Since $3 \times 3$ submatrices can be obtained by removing one row and $n-2$ columns at a time, we have $4 \times \binom{n+1}{3}$ possible matching tensors $T$. However, using simulations we observed that not all of them can be estimated uniquely from point matches. Let $T_{ijk;i'j'k'}$ represent the tensor corresponding to the submatrix with rows $(i, j, k)$ and columns $(i', j', k')$. In the following, we use $2 \times (n-1)$ constraints of the form $T_{x34;12y}, (x = 1, 2;\ y = 3 \cdots n)$ for calibration, i.e. constraints combining the optical center and the first grid, with the other grids. For these tensors, the equation (3) takes the following form: $\sum_{i=1}^{9} C_i^y V_i^y = 0$ and $\sum_{i=1}^{9} C_i^y W_i^y = 0$ for $T_{134;12y}$ and $T_{234;12y}$ respectively. Here, $C_i^y = Q_j^1 Q_k^y$, for appropriate indices $j$, as shown in Table 1.

$V_i^y$ and $W_i^y$ are computed up to scale using least squares. Note that they share some coefficients (e.g. $R_{3,1}^y$), hence they can be estimated up to the same scale factor, $\lambda_y$. We perform this step for $(n-1)$ constraints by choosing $y = 3 \cdots n$. We now combine all the coupled variables

Table 1. Tensors $T_{134;12y}$ and $T_{234;12y}$ for a central camera.

| i | $j$ | $k$ | $V_i^y$ | $W_i^y$ |
|---|---|---|---|---|
| 1 | 1 | 1 | $0$ | $R_{3,1}^y$ |
| 2 | 1 | 2 | $0$ | $R_{3,2}^y$ |
| 3 | 1 | 4 | $0$ | $-O_3 + t_3^y$ |
| 4 | 2 | 1 | $R_{3,1}^y$ | $0$ |
| 5 | 2 | 2 | $R_{3,2}^y$ | $0$ |
| 6 | 2 | 4 | $-O_3 + t_3^y$ | $0$ |
| 7 | 4 | 1 | $-O_2 R_{3,1}^y + O_3 R_{2,1}^y$ | $-O_1 R_{3,1}^y + O_3 R_{1,1}^y$ |
| 8 | 4 | 2 | $-O_2 R_{3,2}^y + O_3 R_{2,2}^y$ | $-O_1 R_{3,2}^y + O_3 R_{1,2}^y$ |
| 9 | 4 | 4 | $-O_2 t_3^y + O_3 t_2^y$ | $-O_1 t_3^y + O_3 t_1^y$ |

contained in the different tensors, to obtain the following system which links the pose variables of all the grids.

$$\begin{bmatrix} H^2_{6\times 2} & J_{6\times 6} & .. & 0_{6\times 6} \\ .. & .. & .. & .. \\ H^n_{6\times 2} & 0_{6\times 6} & .. & J_{6\times 6} \end{bmatrix} \begin{bmatrix} -O_1 \\ -O_2 \\ X^2_{6\times 1} \\ .. \\ X^n_{6\times 1} \end{bmatrix} = \begin{bmatrix} Y^2_{6\times 1} \\ .. \\ Y^n_{6\times 1} \end{bmatrix}, \quad (4)$$

$$H^i = \begin{bmatrix} 0 & V_4^i \\ 0 & V_5^i \\ 0 & V_6^i \\ V_4^i & 0 \\ V_5^i & 0 \\ V_6^i & 0 \end{bmatrix}, J = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$X^i = \begin{bmatrix} \lambda_i O_3 (t_1^i - O_1) \\ \lambda_i O_3 (t_2^i - O_2) \\ \lambda_i O_3 R_{1,1}^i \\ \lambda_i O_3 R_{1,2}^i \\ \lambda_i O_3 R_{2,1}^i \\ \lambda_i O_3 R_{2,2}^i \end{bmatrix}, Y^i = \begin{bmatrix} V_7^i \\ V_8^i \\ V_9^i \\ W_7^i \\ W_8^i \\ W_9^i \end{bmatrix}$$

We rewrite equation (4) as follows:

$$A_{6(n-1)\times(2+6(n-1))} Z_{2+6(n-1)} = Y_{6(n-1)}$$

Since $A$ is of rank $6(n-1)$, we obtain the $(2+6(n-1))$ variables $(Z's)$ up to a linear combination of three vectors. The coefficients of the linear combination are computed using orthogonality constraints on rotation matrices $R^i$. More details are given in [18].

Using the definition of $Z$, it is possible to compute the pose variables uniquely except for a sign ambiguity in $n$ variables: there are two mirror solutions for each grid's pose (they can lie on either side of the optical center). In the case of a pinhole camera we can resolve this ambiguity by applying the constraint that the grids must lie on the same side of the optical center. However this constraint becomes difficult to apply for omnidirectional cameras where the grids essentially get distributed around the center. We apply the following technique. First we arbitrarily select one solution for the first grid's pose. Then we identify the correct location of each of the other grids by minimizing their distance with an already fixed grid, with which it has some overlap. This is easily achieved because we usually collect images in succession and not in a completely random order.

Having determined the pose of grids and the optical center, we now compute projection rays for all pixels that have at least one matching point in one of the grids used here.

## 2.2 Pose Estimation of Additional Grids

We suppose here that a partial calibration of the camera has been performed with the method of the previous section. The calibration is partial because only grids whose projection in the image had some overlap with one of the grids ("the first grid") were used. In order to make the calibration complete, we use the pose estimation technique, described in our earlier work [17], to include additional grids, which do not have any overlap with the first grid, but with some of the others. A 4th degree polynomial equation is solved to compute the pose and the correct solution is identified from the ambiguous ones as given in [9].

## 2.3 Bundle Adjustment

We use bundle adjustment [24] to refine the pose of all grids (except for the first one) and the projection rays. During bundle adjustment, we minimize the generic *ray-point* distance metric [17], i.e. the sum of distances between a grid point and the projection ray of a pixel that has seen that point. This can be applied at any stage of our approach; we apply it after the initial calibration using multiple grids (cf. §2.1), for refining the pose of each additional grid (cf. §2.2), as well as at the end of the whole calibration [18].

# 3. Variants

## 3.1. Non-Central Cameras

In the non-central case, collinearity constraints require 3 or more grid points per pixel, instead of 2 for central cameras (where the optical center, though unknown, is taken into account). We use the same notations as in §2.1. For a non-central camera, we apply the collinearity constraint on the region given by $\cup_{i=3}^{n}(B_1 \cap B_2 \cap B_i)$. Once the poses are computed the calibration is eventually extended to a larger region given by $\cup_{i,j=1\cdots n; i\neq j}(B_i \cap B_j)$.

We now summarize the calibration procedure, analogously to §2.1. We have no optical center here, so do consider the following $4 \times n$ matrix of collinear points:

$$\begin{pmatrix} Q_1^1 & R_{11}^2 Q_1^2 + R_{12}^2 Q_2^2 + t_1^2 Q_4^2 & \cdots \\ Q_2^1 & R_{21}^2 Q_1^2 + R_{22}^2 Q_2^2 + t_2^2 Q_4^2 & \cdots \\ 0 & R_{31}^2 Q_1^2 + R_{32}^2 Q_2^2 + t_3^2 Q_4^2 & \cdots \\ Q_4^1 & Q_4^2 & \cdots \end{pmatrix}$$

Similarly to the central case we can apply the collinearity constraint by equating the determinant of every $3 \times 3$ submatrix to zero. Using simulations we found, as in the central case, that not all of these provide unique solutions. In contrast to the central case, where we used the center and the first grid to build a system linking all the pose variables, we here use the first and second grid to build the system. Thus we have $3 \times (n-2)$ possible tensors, represented by $T_{3jk;12y}, (j,k \in \{1,2,4\},\ y = 3\cdots n)$. As in the central

case, we are able to use these tensors to estimate the poses of all the grids. More details are given in [18].

The next step of the calibration chain, pose estimation and computation of further projection rays, is also slightly different compared to central cameras (cf. §2.2). Here, the calibration region is extended to $C_{k+1} = \cup_{i,j=1\cdots n;i\neq j}(B_i \cap B_j)$, i.e. it contains all pixels that are matched to at least 2 grid points. As for the actual pose estimation, it can be formulated in the same way as for central cameras, but may lead to a set of 8 solutions that does not contain reflected pairs [4, 13, 17]. Disambiguation can be carried out using additional points besides the 3 used for the minimal pose routine.

### 3.2. Slightly Non-Central Cameras

For slightly non-central cameras like fisheye, spherical or hyperbolic catadioptric cameras, we start by running the central version of the generic calibration to obtain an initial partial calibration. Typically we use four or five images simultaneously to calibrate an image region and then use pose estimation to add other images and cover the rest of the image region. Next, we relax the central assumption; projection rays are first computed from grid points, without enforcing them to pass through an optical center. After this, a non-central bundle adjustment is performed [18].

### 3.3. Selecting the Best Camera Model

The non-central calibration algorithm of §3.1, can not be used as such to calibrate a central camera: data (pixel-to-grid correspondences) coming from a central camera, will lead to a higher rank-deficiency in the linear solution of the tensors, causing an incorrect calibration (although residuals will be lower). However, we may, by analyzing the rank of the underlying equation system, detect this problem and maybe even classify the camera as being central and then apply the appropriate calibration algorithm. More generally speaking, this is a model selection problem, and the rank-analysis or any other solution will allow to build a truly complete black box calibration system.

To this end, we have to take into account a few intermediate camera models that may be encountered in practice. One such case is the class of cameras for which there exists a single line that cuts all projection rays (we call them *axial cameras*). Examples are the classical two-camera stereo systems (the mentioned line is the baseline joining the two optical centers) and certain non-central catadioptric cameras, e.g. all catadioptric cameras with a spherical mirror. A yet more special class of cameras are so-called crossed-slits cameras [5], which encompass pushbroom cameras [8]. We are currently specializing our calibration approach to these additional general classes of camera types. Overall, it seems that these 4 classes (central, axial, crossed-slits, fully non-central) and their associated calibration algorithms, maybe with a few additional classes, should be sufficient to calibrate most cameras.

Besides considering these general camera types, we may also discuss the choice between parametric and non-parametric models for a given camera. Generic calibration not only allows to calibrate any camera system by treating it as a black box, it also provides the ability to easily obtain a parametric calibration once the model for the camera is known. Every parametric calibration will just be a model-fitting problem, which can be solved as a non-linear optimization problem starting with the good initial solution obtained using generic calibration.

## 4. Practical Issues

First, we found that grids with circular targets, using point spread function to compute the centers, provide stable calibration compared to checkerboard patterns.

Secondly, the usage of grids with very different orientations and positions is important for stable calibration. One way to easily achieve this is to use calibration grids of different sizes and to put them at different distances from the camera (together with sufficient orientation differences).

Thirdly, by using a combination of *local* 4-point homography based prediction, local collinearity and orthogonality constraints, we start from four features (circular targets or corners), located at the corners of a square, and incrementally extend the matching of image features to grid coordinates along all directions. This approach worked automatically for all pinhole images as well as for several fisheye and catadioptric images. However we also had to use manual input for some images.

The last issue is concerned with a required interpolation process: for every grid point in the first image we compute the interpolated points in the other grids' coordinate systems (since for other grids, the extracted targets or corners do not lie on the same pixels in general). To take care of the noise we impose collinearity constraints (globally for central cameras and locally for non-central cameras) during interpolation process for the originally collinear corners in the calibration grids [18]. This improved the numerical stability of the results significantly.

## 5. Experiments and Results

We have calibrated a wide variety of cameras (both central and non-central) as shown in Table 2. Results are first discussed for several "slightly non-central" cameras, and then for a multi-camera system.

**Slightly non-central cameras: central vs. non-central models.** For three cameras (a fisheye, a hyperbolic and a spherical catadioptric system, see sample images in Figure 3), we applied both, central calibration and the procedure explained in § 3.2, going from central to non-central.

Table 2. Bundle adjustment statistics for different cameras. (C) and (NC) refer to central and non-central calibration respectively, and 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).

| Camera | Images | Rays | Points | RMS |
|---|---|---|---|---|
| Pinhole (C) | 3 | 217 | 651 | 0.04 |
| Fisheye (C) | 23 | 508 | 2314 | 0.12 |
| (NC) | 23 | 342 | 1712 | 0.10 |
| Sphere (C) | 24 | 380 | 1441 | 2.94 |
| (NC) | 24 | 447 | 1726 | 0.37 |
| Hyperbolic (C) | 24 | 293 | 1020 | 0.40 |
| (NC) | 24 | 190 | 821 | 0.34 |
| Multi-Cam (NC) | 3 | 1156 | 3468 | 0.69 |
| Eye+Pinhole (C) | 3 | 29 | 57 | 0.98 |

Table 3. RMS error for circle fits to grid points, for turntable sequences (see text).

| Camera | Grids | Central | Non-Central |
|---|---|---|---|
| Fisheye | 14 | 0.64 | 0.49 |
| Spherical | 19 | 2.40 | 1.60 |
| Hyperbolic | 12 | 0.81 | 1.17 |

Table 2 shows that the bundle adjustment's residual errors for central and non-central calibration, are very close to one another for the fisheye and hyperbolic catadioptric cameras. This suggests that for the cameras used in the experiments, the central model is appropriate. As for the spherical catadioptric camera, the non-central model has a significantly lower residual, which may suggest that a non-central model is better here.

To further investigate this issue we performed another evaluation. A calibration grid was put on a turntable, and images were acquired for different turntable positions. We are thus able to quantitatively evaluate the calibration, by measuring how close the recovered grid pose corresponds to a turntable sequence. Individual grid points move on a circle in 3D; we thus compute a least squares circle fit to the 3D positions given by the estimated grid pose. At the bottom of Figure 3, recovered grid poses are shown, as well as a circle fit to the positions of one grid point. Table 3 shows the RMS errors of circle fits (again, relative to scene size, and given in percent). We note that the non-central model provides a significantly better reconstruction than the central one for the spherical catadioptric camera, which thus confirms the above observation. For the fisheye, the non-central calibration also performs better, but not as significantly. As for the hyperbolic catadioptric camera, the central model gives a better reconstruction though. This can probably be explained as follows. Inspite potential imprecisions in the camera setup, the camera seems to be sufficiently close to a central one, so that the non-central model leads to overfitting. Consequently, although the bundle adjustment's residual is lower than for the central model (which always has to be the case), it gives "predictions" (here, pose or motion estimation) which are unreliable.

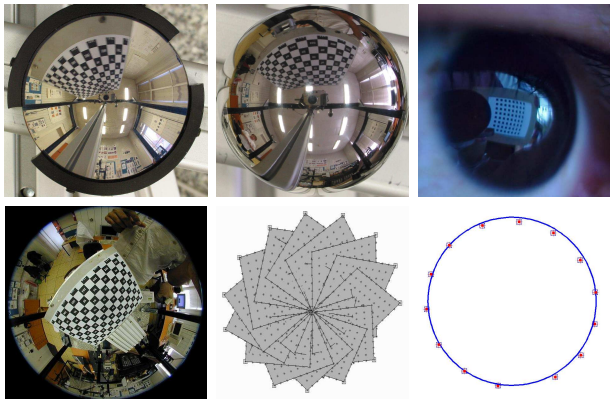


Figure 3. Top: sample images for hyperbolic (left), spherical (middle) and eye based catadioptric cameras (right). Bottom: fisyeye image (left), pose of calibration grids used to calibrate the fisheye (middle) and a least squares circle fit to the estimated positions of one grid point (right).

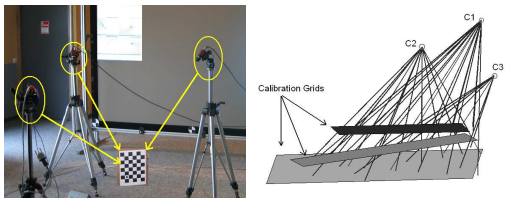



Figure 4. Multi-camera setup consisting of 3 cameras (left). Recovered projection rays and grid poses (right).

**Calibration of a multi-camera system.** A multi-camera network can be considered as a single generic imaging system. As shown in Figure 4 (left), we used a system of three (approximately pinhole) cameras to capture three images each of a calibration grid. We virtually concatenated the images from the individual cameras and computed all projection rays and the three grid poses in a single reference frame (see Figure 4 (right)), using the non-central algorithm described in § 3.1.

In order to evaluate the calibration, we compared results with those obtained by plane-based calibration [21, 27], that used the knowledge that the three cameras are pinholes. In both, our multi-camera 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 4). As for translation, we measure the

Table 4. Evaluation of non-central multi-camera calibration relative to plane-based calibration. See text for more details.

| Camera | $R_2$ | $R_3$ | $t_2$ | $t_3$ | Center |
|---|---|---|---|---|---|
| 1 | 0.0117 | 0.0359 | 0.56 | 3.04 | 2.78 |
| 2 | 0.0149 | 0.0085 | 0.44 | 2.80 | 2.17 |
| 3 | 0.0088 | 0.0249 | 0.53 | 2.59 | 1.16 |

distance between the estimated 3D positions of the grids' centers of gravity (columns $t_i$ in Table 4) expressed in percent, relative to the scene size. Here, plane-based calibration is done separately for each camera, leading to the three rows of Table 4.

From the non-central multi-camera calibration, we also estimate the positions of the three optical centers, by clustering the projection rays and computing least squares point fits to them. The column "Center" of Table 4 shows the distances between optical centers (expressed in percent and relative to the scene size) computed using this approach and plane-based calibration. The discrepancies are low, suggesting that the non-central calibration of a multi-camera setup is indeed feasible.

Another experiment we carried out was to calibrate a small region of the exotic catadioptric system formed with an eye as mirror, cf. an image in Figure 3 and bundle adjustment statistics in Table 2.

## 6. Summary and Conclusions

We have proposed a non-parametric, generic calibration approach and shown its feasibility by calibrating a wide variety of cameras. One of the important issues is in the identification of appropriate models, central or non-central, for slightly non-central cameras. For understanding complex cameras or mirror surfaces, where mathematical modeling might be more demanding, generic calibration can be used as a black box tool to first obtain the projection rays. The nature of these projection rays can be experimented further to identify the right parametric model.

**Acknowledgments.** We want to thank Tomaš Pajdla, Branislav Mičušik, Bertrand Holveck and Thomas Bonfort for their help in the experiments. This work was partially supported by the Multi University Research Initiative (MURI) grant by Army Research Office under contract DAA19-00-1-0352 and the NSF grant ACI-0222900.

## 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] G. Champleboux, S. Lavallée, P. Sautot and P. Cinquin. Accurate Calibration of Cameras and Range Imaging Sensors: the NPBS Method. *ICRA*, 1992.

[4] C.-S. Chen and W.-Y. Chang. On Pose Recovery for Generalized Visual Sensors. *PAMI*, 2004.

[5] D. Feldman, T. Pajdla and D. Weinshall. On the Epipolar Geometry of the Crossed-Slits Projection. *ICCV*, 2003.

[6] K.D. Gremban, C.E. Thorpe, and T. Kanade. Geometric Camera Calibration using Systems of Linear Equations. *ICRA*, 1988.

[7] M.D. Grossberg and S.K. Nayar. A General Imaging Model and a Method for Finding its Parameters. *ICCV*, 2001.

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

[9] R.M. Haralick, C.N. Lee, K. Ottenberg, and M. Nolle. Review and Analysis of Solutions of the Three Point Perspective Pose Estimation Problem. *IJCV*, 1994.

[10] R.A. Hicks and R. Bajcsy. Catadioptric Sensors that Approximate Wide-angle Perspective Projections. *CVPR*, 2000.

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

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

[13] D. Nistér. A Minimal Solution to the Generalized 3-Point Pose Problem. *CVPR*, 2004.

[14] T. Pajdla. Stereo with Oblique Cameras. *IJCV*, 2002.

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

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

[17] S. Ramalingam, S.K. Lodha, and P. Sturm. A Generic Cross-Camera Structure-from-Motion Analysis. *Omnivis*, 2004.

[18] S. Ramalingam, P. Sturm, and S.K. Lodha. Theory and Experiments towards Complete Generic Calibration. *Research Report*, INRIA, April, 2005.

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

[20] H.Y. Shum and L.W. He. Rendering with Concentric Mosaics. *SIGGRAPH*, 1999.

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

[22] P. Sturm and S. Ramalingam. A Generic Concept for Camera Calibration. *ECCV*, 2004.

[23] R. Swaminathan, M.D. Grossberg, and S.K. Nayar. A Perspective on Distortions. *CVPR*, 2003.

[24] B. Triggs, P. McLauchlan, R. Hartley, and A. Fitzgibbon. Bundle Adjustment: A Modern Synthesis. *Workshop on Vision Algorithms: Theory and Practice*, 2000.

[25] Y. Wexler, A.W. Fitzgibbon, and A. Zisserman. Learning Epipolar Geometry from Image Sequences. *CVPR*, 2003.

[26] D.N. Wood, A. Finkelstein, J.F. Hughes, C.E. Thayer, and D.H. Salesin. Multiperspective Panoramas for Cel Animation. *SIGGRAPH*, 1997.

[27] Z. Zhang. A Flexible New Technique for Camera Calibration. *PAMI*, 2000.