

# On the Equivalence of Moving Entrance Pupil and Radial Distortion for Camera Calibration

Avinash Kumar and Narendra Ahuja

Department of Electrical and Computer Engineering

University of Illinois at Urbana-Champaign, Urbana, 61801, IL, USA

{avinash,n-ahuja}@illinois.edu

## Abstract

*Radial distortion for ordinary (non-fisheye) camera lenses has traditionally been modeled as an infinite series function of radial location of an image pixel from the image center. While there has been enough empirical evidence to show that such a model is accurate and sufficient for radial distortion calibration, there has not been much analysis on the geometric/physical understanding of radial distortion from a camera calibration perspective. In this paper, we show using a thick-lens imaging model, that the variation of entrance pupil location as a function of incident image ray angle is directly responsible for radial distortion in captured images. Thus, unlike as proposed in the current state-of-the-art in camera calibration, radial distortion and entrance pupil movement are equivalent and need not be modeled together.*

*By modeling only entrance pupil motion instead of radial distortion, we achieve two main benefits; first, we obtain comparable if not better pixel re-projection error than traditional methods; second, and more importantly, we directly back-project a radially distorted image pixel along the true image ray which formed it. Using a thick-lens setting, we show that such a back-projection is more accurate than the two-step method of undistorting an image pixel and then back-projecting it. We have applied this calibration method to the problem of generative depth-from-focus using focal stack to get accurate depth estimates.*

## 1. Introduction

Camera calibration entails estimation of intrinsic and extrinsic parameters of a camera given a set of known world points and their measured image coordinates on the image plane [3, 13, 14, 7, 16]. The intrinsic parameters model the geometric properties of the image sensor and the extrinsic properties model the pose of the camera in a known coordinate system.

Camera calibration depends on the imaging model that governs the camera optics. In this paper, we focus on or-

dinary (non-fisheye) camera lenses whose image formation can be described as ideal perspective projection followed by small deviations/distortions. These distortions cause straight lines in the scene to be imaged as curved lines. In calibration, image distortion has been modeled as a combination of radial distortion [3] to model radially outward movement of ideal image points and decentering distortion to model keystone like distortions owing to small lens-sensor misalignment [14, 4]. Both of these distortion models are in the form of an infinite series function of ideal image coordinates [14, 16, 7]. Since, they are not physically motivated, these models are only approximate in nature. Previously, Kumar and Ahuja [10] have proposed a physical model for decentering distortion via a rotation matrix to encode the lens-sensor tilt. In this paper, we show that the physical motion of the entrance pupil as a function of the incident ray angle is the cause of radial distortion in images. Based on this analysis, we propose a new camera calibration method which features a pupil-centric imaging model [10] (more accurate than traditional pin-hole model), rotation-based image sensor misalignment [10, 5] (to model decentering effects) and moving entrance pupil model (to model radial distortion) to achieve lower pixel re-projection errors. Moreover, our calibration allows for accurate back-projection of image rays directly from distorted image pixels. Our analysis shows that the combined modeling of entrance pupil motion and radial distortion [14] as done in the current state-of-the-art in camera calibration [5] is redundant. Thus, our main contributions in this paper are:

1. In a pupil-centric imaging framework, we show that the physical motion of entrance pupil along the optic axis causes radial bending of image rays thus resulting in radial distortion of image pixels (Sec. 2).
2. We propose a new pupil-centric calibration algorithm based on these ideas and obtain less pixel re-projection error compared to prior calibration methods (Sec. 3) and show accurate pixel ray back-projection (Sec. 5).
3. We present a categorization of various prior calibration algorithms based on their underlying assumptions and

show how the proposed model is new and differs from all of them (Sec. 4).

4. We present a generative depth-from-focus algorithm which depends on accurate ray back-projection and uses the proposed calibration method to compute scene depth (Sec. 5).

In Sec. 6, we compare our calibration results with those obtained by prior methods and show that we get least re-projection error without increasing the parameterization of the calibration problem. Finally, in Sec. 7, we discuss the limitation of our calibration algorithm with respect to image distortion and propose alternate approximations to overcome them.

## 2. The New Insight on Radial Distortion Modeling

In this section, we first present definitions of some lens parameters useful to explain our radial distortion model. Then, we present the traditional radial distortion model which assumes pin-hole imaging. This is followed by the proposed entrance pupil motion based modeling of radial distortion for thin and thick lenses.

**Lens parameters:** See Fig. 1. A typical lens consists of a system of lenses with a physical aperture stop somewhere between these lenses. This stop physically limits the amount of light and controls the depth of field captured in focus by the lens. The entrance pupil ( $E_n$ ) is defined as the virtual image of the aperture stop as seen from the front of the lens. Similarly, the exit pupil ( $E_x$ ) is the image of the aperture stop as seen from behind the lens. For a given image pixel, the corresponding image ray forming that pixel, typically referred to as the chief ray, must physically pass through the aperture stop. Thus, it also appears to pass through the center of the entrance and exit pupil. In camera calibration, most of the prior works assume these points to be fixed, except [5], which models their motion. In this work, we will assume that entrance pupil moves.

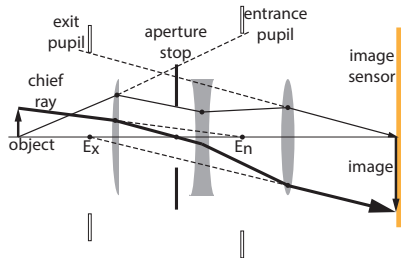


Figure 1. Typical lens design: entrance, exit pupil, aperture stop.

**Conventional radial distortion modeling:** Let an ideal perspective projection of a world point on the image plane be  $P_p = (x_p, y_p)$ . Due to distortion, this point actually appears at location  $P_d = (x_d, y_d)$ . Assuming no decentering

of lens-sensor, these two points can be related by only radial distortion parameters ( $k_1, k_2$ ) as:

$$\begin{bmatrix} x_d \\ y_d \end{bmatrix} = (1 + k_1 \cdot r_p^2 + k_2 \cdot r_p^4) \begin{bmatrix} x_u \\ y_u \end{bmatrix} \quad (1)$$

where  $r_p = \sqrt{x_u^2 + y_u^2}$ . This is conventional radial distortion modeling as in [3, 14, 7, 16].

**Thin-lens: stop location and radial distortion:** Consider a thin-lens setting as shown in Fig. 2(a, b, c) where the chief ray forms the image, while the stop location varies. We observe that depending upon the stop location, the chief

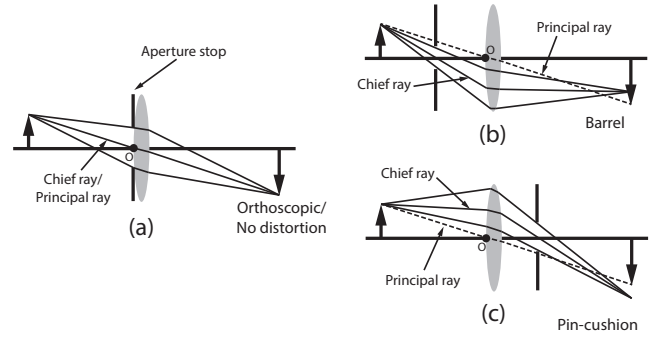


Figure 2. (a) No-distortion: aperture stop coincides with the principal point  $O$ . (b,c) Barrel and pin-cushion distortion due to change in stop location.

ray bends as it does not pass through the center of the lens. It is known that any ray passing through the lens center remain undeviated [6], thus forming ideal perspective image. Thus, we can conclude that changing the stop location causes the resulting image to distort on the image plane.

**Thick-lens: stop location and radial distortion** Now, we model the lens as being thick-lens and extend the above analysis to relate radial distortion and stop location. A thick-lens model follows a pupil-centric imaging geometry as shown in Fig. 3. which is parameterized by a set

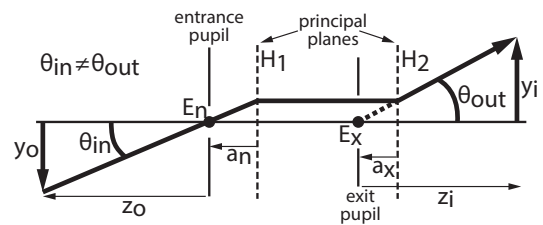


Figure 3. Pupil-centric imaging model for thick lenses.

of known lens parameters, namely entrance pupil  $E_n$ , exit pupil  $E_x$  and the front and back principal planes  $H_1, H_2$ . The chief ray passes through  $E_n$  and exits through  $E_x$ , making an angle of  $\theta_{in}$  at  $E_n$  and  $\theta_{out}$  at  $E_x$ . It was shown

in [1], that  $\theta_{in}$  and  $\theta_{out}$  are related as

$$\tan(\theta_{in}) = \frac{F}{F - a_x} \tan(\theta_{out}) \quad (2)$$

where  $F$  is the optical focal length of the system. If  $y_o$  and  $y_i$  are object and image heights respectively and  $z_o$  and  $z_i$  are object and image distances from  $E_n$  and  $E_x$  respectively, then we have  $\tan(\theta_{in}) = \frac{y_o}{z_o}$  and  $\tan(\theta_{out}) = \frac{y_i}{z_i}$ . The transverse magnification  $M_T$  [6] of the system can then be computed from Eq. 2 as:

$$M_T = \frac{y_i}{y_o} = \frac{F - a_x}{F} * \frac{z_i}{z_o} \quad (3)$$

If  $z_o$  and  $z_i$  are fixed, then  $M_T$  is fixed as  $a_x$  and  $F$  are fixed. Thus, we can conclude that for fixed lens parameters there is a constant magnification factor relating all object points on a fixed plane at  $z_o$  and the captured image points. But, we know from observed images captured from a thick lens, that magnification is not constant across the image plane, rather it either increases or decreases as the image points move far from the center of the image. Or in other words, the magnification  $M_T$  is radially varying causing observed radial distortion of image points. Thus, we can give the following hypothesis.

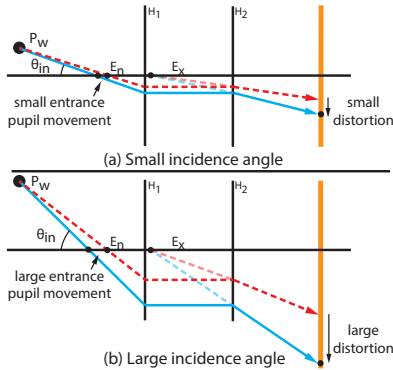


Figure 4. Angle of incident light ray moves the entrance pupil leading to different amounts of radial distortion: (a) small incident angle (b) large incident angle.

**Our hypothesis: moving entrance pupil causes radial distortion in pupil-centric setting** The only way to have varying magnification/radial distortion over the image plane is to assume that all of the ray geometries shown in Fig. 2(a,b,c) occur simultaneously. This would imply that there are multiple aperture stop locations in the imaging system at the same time. Since, this is physically impossible, we hypothesize that in the case of a thick lens, the entrance pupil varies its location monotonically depending on the angle  $\theta_{in}$  of incidence of the incoming ray. Such phenomenon

has previously been observed for fish-eye lenses by Genery [5]. This allows us to explain for small angles of incidence (Fig. 4(a)), we have ideal perspective projection (similar to Fig 2(a)), agreeing with the observation that there is almost no radial distortion at regions near the image center. But as the incident ray angle  $\theta_{in}$  increases (Fig. 4(b)), the entrance pupil location changes and the imaging system behaves similar to Fig. 2(b) or (c). This confirms to the observation that radial distortion is higher towards the periphery of the image. Thus, we can conclude that the movement of the entrance pupil as a function of incident ray angle  $\theta_{in}$  is a predominant geometric reason for the occurrence of radial distortion on the image plane. Next, based on our hypothesis, we present our new camera calibration model.

### 3. Proposed Camera Calibration Modeling

Our camera calibration model assumes pupil-centric imaging with moving entrance pupil and non-frontal sensor model [10] as shown in Fig. 5, where a known world point location  $P_w$  is imaged to pixel  $P_i$  based on current calibration parameters and  $\hat{P}_i$  is the actual measured image coordinate. We define the following coordinates systems:

- **world coordinate system ( $C_w$ ):** The known world points are defined in this coordinate system e.g. a checkerboard.
- **entrance pupil coordinate system ( $C_{E_n}$ ):** The origin of this coordinate system lies at the entrance pupil  $E_n$  of the imaging system. Since, in our case  $E_n$  is assumed to be moving with respect to the incident image ray (Sec. 2), we define the origin of  $C_{E_n}$  to correspond to the location where the incidence rays with  $\theta_{in} \approx 0$  intersect the optic axis. This location can typically be obtained from the lens data sheet as a signed distance from the front principal plane.
- **sensor coordinate system ( $C_s$ ):** The  $xy$  plane of this coordinate system lies on the image sensor with origin at the intersection of optic axis and the sensor. This intersection point is the center of radial distortion (CoD).
- **image coordinate system ( $C_i$ ):** The measured pixel values are described in this coordinate system.

The distances are measured in pixels in  $C_i$  and in metric (e.g. mm) in  $C_w, C_{E_n}, C_s$ . See Fig. 5. Let the world point be  $P_w = (X, Y, Z)$  and the measured image point be  $\hat{P}_i = (\hat{I}, \hat{J})$ . Let the signed distance of  $E_n$  from  $H_1$  be  $a_n$  and of  $E_x$  from  $H_2$  be  $a_x$ . Ignoring noise, we have the following transformations relating  $P_w$  and  $\hat{P}_i$ .

**Transformation from  $C_w$  to  $C_{E_n}$ :** Let  $S = \{s_{ij} : 1 \leq (i, j) \leq 3\}$  be the rotation matrix and  $T = (t_x, t_y, t_z)$  be the translation between these two coordinate systems. Then,  $P_w$  can be expressed as  $P_l = (x_l, y_l, z_l)$  in  $C_{E_n}$  as:

$$\begin{bmatrix} x_l \\ y_l \\ z_l \end{bmatrix} = \begin{bmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ s_{31} & s_{32} & s_{33} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad (4)$$

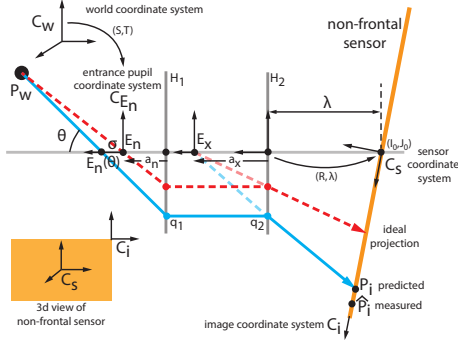


Figure 5. Geometry of proposed calibration model.

**Incorporating entrance pupil movement:** Let the ray from  $P_w$  be incident on the optical axis at  $E_n(\theta)$  at an angle of  $\theta$ . Let the distance between  $E_n$  and  $E_n(\theta)$  be  $\sigma(\theta)$ , where we use the model for  $\sigma(\theta)$  from [5] as:

$$\sigma(\theta) = \left( \frac{\theta}{\sin(\theta)} - 1 \right) (\epsilon_1 + \epsilon_2 \theta^2 + \dots) \quad (5)$$

Here  $(\epsilon_1, \epsilon_2)$  are the pupil movement parameters. As the distance moved by the pupil  $\sigma(\theta)$  and  $\theta$  are dependent on each other they cannot be determined independently. But, a simple trigonometric equation can be derived in terms of  $P_i$  and  $\theta$  [5] as:

$$z_l \sin(\theta) - \sqrt{x_l^2 + y_l^2} \cos(\theta) - (\theta - \sin(\theta))(\epsilon_1 + \epsilon_2 \theta^2) = 0$$

and can be solved iteratively for  $\theta$  using Newton-Raphson method using initial value of  $\theta$  as  $\arctan\left(\frac{\sqrt{x_l^2 + y_l^2}}{z_l}\right)$  and current estimates of calibration parameters  $S, T, \epsilon_1, \epsilon_2$ . An example plot of the variation of entrance pupil center as incident ray angle  $\theta$  changes is shown in Fig. 6 for the lens used in our experiments. As can be seen, the default  $E_n$  provided by the manufacturer has been calibrated for incident rays with  $\theta \approx 50$  degrees.

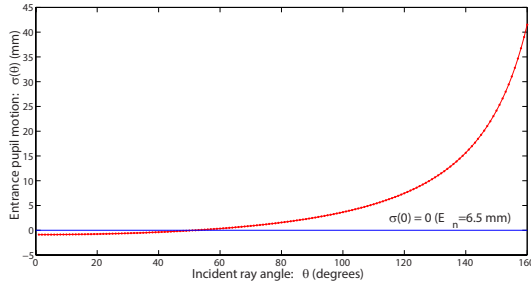


Figure 6. Moving entrance pupil vs angle of incidence ray. The plot is based on  $(\epsilon_1 = -5.304, \epsilon_2 = 6.474)$  computed from calibration results from our proposed method as shown in the last column of Table. 1.

**Computing incident ray intersection with  $H_1$ :** Given the computed ray direction  $\theta$ , the incident ray from  $P_w$  intersects the front principal plane  $H_1$  at some location  $q_1$ . Using coordinate geometry, the coordinates of  $q_1$  are:

$$q_1 = \frac{-(a_n + \sigma(\theta))}{z_l - \sigma(\theta)} \begin{bmatrix} x_l \\ y_l \end{bmatrix} \quad (6)$$

As ray incidence on  $H_1$  and  $H_2$  are same, we have  $q_1 = q_2$ .

**Computing the exitance ray from  $E_x$  to  $q_2$ :** As the exit ray must appear to come from  $E_x$ , its intersection with the non-frontal sensor plane can be computed. We assume that the image sensor is rotated with respect to  $H_2$  by a two parameter rotation matrix  $R(\alpha, \beta)$  [10]. We also assume that the origin of the non-frontal sensor is located at a distance of  $\lambda$  from  $H_2$ . We have  $E_x = (0, 0, a_x)$ . Now, both  $E_x$  and  $q_2$  can be obtained in terms of  $C_s$  as:

$$E_x^{C_s} = R \begin{bmatrix} 0 & 0 & a_x + \lambda \end{bmatrix}^t \quad (7)$$

$$q_2^{C_s} = R \begin{bmatrix} q_2(x) & q_2(y) & \lambda \end{bmatrix}^t \quad (8)$$

where,  $[\ ]^t$  denotes transpose. Let the intersection of this ray with the image sensor be  $P_s = (x_s, y_s)$ .

**Obtaining image coordinates:** If the pixel sizes are  $s_x$  and  $s_y$  and the origin of  $C_s$  is at  $(I_0, J_0)$  on the image, then  $P_s$  can be transformed to pixel coordinates  $P_i = (I, J)$  as

$$\begin{bmatrix} I \\ J \end{bmatrix} = \begin{bmatrix} \frac{x_s}{s_x} - I_0 \\ \frac{y_s}{s_y} - J_0 \end{bmatrix} \quad (9)$$

We refer to  $(I_0, J_0)$  as the center of distortion (CoD). Thus  $P_i$  is obtained in terms of 14 calibration parameters  $U$  as:

$$U = \{ \underbrace{S, T}_{\text{extrinsic}(6)}, \underbrace{R(\alpha, \beta), s_x, \lambda, I_0, J_0, \epsilon_1, \epsilon_2}_{\text{intrinsic}(8)} \} \quad (10)$$

We note that  $s_y$  is assumed to be known as it decides the scale of the calibration estimates [13, 14, 16].

**Linear and nonlinear optimization:** The calibration is done in two stages (1) initial linear estimation and (2) final nonlinear refinement using the estimated parameters from linear estimation. For stage (1), we assume that the entrance pupil is fixed and used the analytical technique in [10]. The final nonlinear estimation is done using Levenberg-Marquardt optimization [7] by minimizing the pixel re-projection error over  $N$  world-image point correspondences:

$$U^* = \arg \min_U \sum_{n=1}^N \|\hat{P}_i^n - P_i^n(U)\|_2^2 \quad (11)$$

where  $U^*$  is the final optimal results. The results in the columns of Table. 1 for various calibration techniques correspond to  $U^*$  from respective methods.



## 4. Comparison with Other Calibration Algorithms

In this section, we describe a subset of existing state-of-the-art in camera calibration with which we would be comparing our proposed model's calibration accuracy in the results section (Sec. 6). These methods vary with respect to each other in terms of the imaging model, the distortion model and the orientation of the image sensor. We next explain these criteria and various models which fall under these criteria.

**Imaging model :** The imaging model describes the image formation from the world point to the image point. There are two types of imaging model being employed in camera calibration. The basic thin-lens model assumes that the incident and exiting rays responsible for image formation by the optical system are principal rays which pass through the optic center of the system of lenses and are parallel to each other [3, 13, 14, 7, 16]. The second model called as pupil-centric model assumes that the image rays responsible for image formation are the chief rays which enter the imaging system at the entrance pupil and appear to exit from the exit pupil [6, 1]. In this model, the location of the entrance pupil can either be assumed to be fixed [10] or it can be assumed to be moving [5].

**Distortion model:** Real imaging systems behave far from ideal perspective projection and are often accompanied by some amount of visible distortion on the image plane where straight lines in real world are imaged as curves. The distortion can be modeled as a combination of radial and decentering distortion [3, 13, 14, 7, 16]. The model is basically an infinite series function of the ideal image points.

**Orientation of the image sensor:** Many times the imaging surface may not be normal to the physical optic axis of the lens system due to manufacturing limitations or sometimes to achieve special effects, e.g. tilt-shift effect. Traditionally, it has been assumed that there exists an effective optic axis which is normal to the image sensor plane. This is referred to as a frontal sensor [14, 7] model. Recently, it has been proposed that calibration can be designed about the physical optic axis by assuming that the sensor is non-frontal with respect to the lens ( $H_2$ ) plane. The non-frontalness can be modeled as a two parameter rotation matrix relating the lens plane and the image sensor plane. This is called as non-frontal sensor modeling [5, 10].

Based on the above criteria, we can classify many prior camera calibration techniques into the following three categories, shown graphically in Fig. 7.

**Category 1:** See Fig. 7(a) for the calibration model in this category. A number of existing calibration methods fall in this category including those proposed in Weng [14], Heikkila [7], Zhang [16]. In this category, imaging is assumed to be thin-lens, the sensor is frontal and image distortion is modeled as a combination of explicit radial and

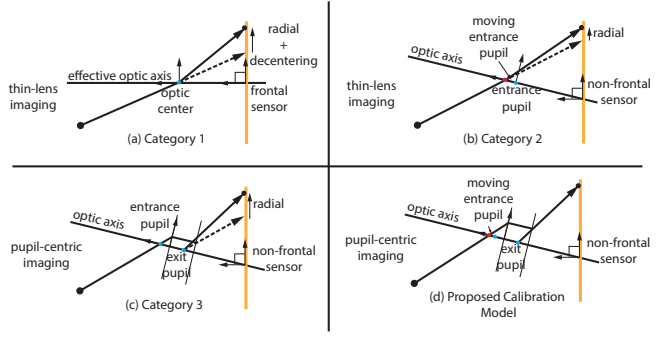


Figure 7. Various calibration models which have been used in literature. (a) Category 1/Classical Method : thin-lens imaging, frontal sensor, radial and decentering distortion model. (b) Category 2 : thin-lens imaging, moving entrance pupil, non-frontal sensor, radial distortion model. (c) Category 3 : pupil-centric imaging, fixed entrance pupil, non-frontal sensor radial distortion. (d) Proposed Calibration Model: pupil-centric imaging, moving entrance pupil, non-frontal sensor.

decentering distortion.

**Category 2** (Gennery [5]): See Fig. 7(b) for the complete model. Here the imaging model is thin-lens and the image forming incident and the exiting rays are incident at the entrance pupil location instead of the optic center. But, the entrance pupil is not fixed and is assumed to be moving depending on the incident ray angle. The image sensor is assumed to be non-frontal, i.e. calibration is modeled about the physical optic axis. Any observed image distortion is modeled as explicitly being radial about the physical optic axis.

**Category 3** (Kumar [10]): See Fig. 7(c) for this model. The imaging model here is pupil-centric with incident ray entering the lens system at the entrance pupil and exiting the system at the exit pupil. The entrance pupil is assumed to be fixed. The image sensor is assumed to be non-frontal and it is shown that sensor non-frontalness compensates for decentering distortion adjustment typically done in Category 1 techniques. The only distortion that is modeled is radial distortion using traditional infinite series formulation. Compared to all these methods, our proposed method does not fall in any of these categories as our imaging model is pupil-centric, assumes moving entrance pupil and the image sensor is assumed to be non-frontal and we don't propose to explicitly model the image distortion based on our analysis on equivalence of moving entrance pupil and observed radial distortion. We also incorporate non-frontal sensor model as it physically corresponds to decentering distortion effects and is more robust for large sensor tilts.

## 5. Ray Back-Projection and Depth From Focus

Since the proposed calibration model does an accurate modeling of forward image projection, its easy to back

project an image pixel accurately (see the blue solid line in Fig. 5). Comparatively, the traditional radial distortion model first corrects for image distortion on the image plane and then back-projects it. Lets assume in Fig. 5, that the undistorted point corresponds to ideal image coordinates. The back-projected ray (see the red dotted line in Fig. 5) is altogether different from the actual image forming blue solid ray. Any vision algorithm depending on ray back-projection analysis might get effected due to this discrepancy. Next, we present a depth from focus algorithm which depends on accurate ray back-projection and where we have applied the results of our proposed camera calibration.

Our depth from focus algorithm takes a focal stack as input. A focal stack is a set of registered images which focus at different scene depths in each image such that each scene point is in focus (imaged sharply) in at-least one image frame. Given this focal stack, an all-focused image of the scene can be computed using techniques from [9, 15]. We have the image formation model  $y = k \star x$ , where  $y$  is the observed image,  $k$  is the depth dependent blur kernel and  $x$  is the ideal sharp image. For our case,  $y$  is the focal stack and  $x$  is the all-focused image which we have already computed and the goal is to compute scene depth encoded in  $k$ . Our generative algorithm can be described as:

- Input: camera calibration parameters  $U^*$  (Eq. 11), all-focused image  $I_f$ , focal stack  $FS$ .
- Repeat for all pixel locations  $p$  in  $I_f$ , denoted as  $I_f(p)$ 
  - Back-project pixel  $p$  to an image ray  $\vec{p}$  using  $U^*$ .
  - For discrete depths  $d$  along  $\vec{p}$ , obtain the hypothesized 3D location of  $p$  as  $P(x_d, y_d, d)$ , where  $(x_d, y_d)$  can be obtained from ray geometry of  $\vec{p}$ .
  - Using  $U^*$ , forward project  $P(x_d, y_d, d)$  onto all the other images in the  $FS$  and compute the blur kernels  $k'$  for each image in  $FS$ .
  - Synthetically generate blurred image windows  $y = k' \star I_f$  around  $p$  in all other  $FS$  images.
  - Compare the sum of squared pixel wise error between synthesized and observed images.
  - Select the depth  $d$  which gives minimum error.
- Output: 3D scene depth

The computed depth using our depth from focus algorithm and the calibration estimates  $U^*$  using the proposed calibration method are showed on a set of different scenes is shown in Fig. 8 and Fig. 9. Ray back projection has also been used for SLAM in [12].

## 6. Pixel Re-projection Results

**Calibration data:** The calibration data consists of a precisely constructed glass checkerboard with  $5 \times 5$  mm squares. Since the checkerboard is transparent, it is back lit to generate white and black squares on the captured calibration images. The checkerboard is fixed at a location and a camera with a tilted image sensor is used to capture a set of 5 images of the checkerboard from different viewpoints. A

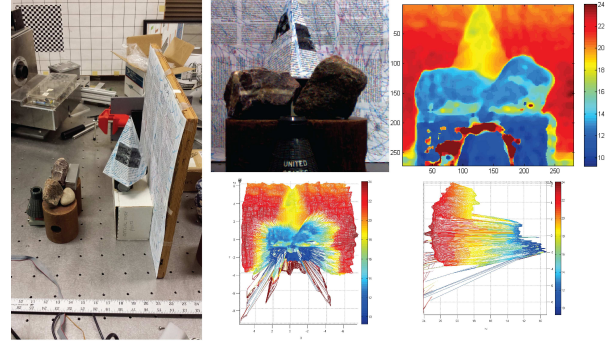


Figure 8. Rocks dataset: (left) shows the setting of various geometric objects in the scene. The objects consist of a flat planar board, a prism object and more generic shaped objects in the form of rocks; (middle top) shows the omnifocus image of the scene; (right top) shows the medial filtered ( $11 \times 11$ ) 2D depth map; (middle and right bottom) shows two 3D views of this dataset. For non-paraxial image rays, Petzval field curvature distortion [6] causes planes to focus on a curved surface. We observe this as the curved reconstruction of the planar board.

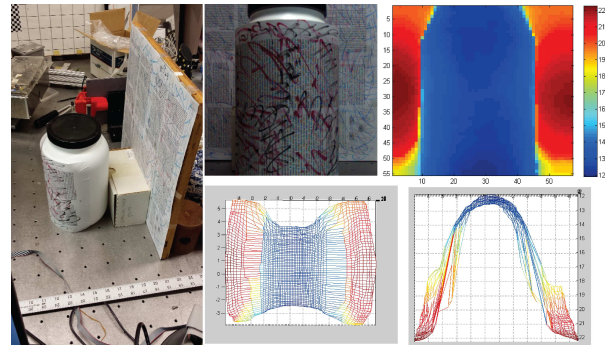


Figure 9. Cylinder dataset: (left) shows a cylindrical container and a flat board as the geometric objects in the scene; (middle top) shows the omnifocus image; (top right) shows the median filtered ( $11 \times 11$ ) 2D depth map of the cylinder; (middle and right bottom) shows two 3D view of the cylinder.

tilted sensor camera is useful in validating the non-frontal modeling in [5, 10]. The corners from the checkerboard images are detected using MATLAB Bouguet’s toolbox [2]. The accuracy of corner detection is separately calculated using the method of [11] and is found to be  $\approx 0.011$  pixels.

**Camera specifications:** We use an AVT Marlin F033C camera with a custom made image sensor which has been slightly tilted by about 3 degrees. This camera is fitted with a Cinegon 1.4/8.2 mm lens. The data sheet [8] of the lens provides the pupil-centric parameters of the camera. In Fig. 10, we show the various parameters provided by the manufacturer. Out of these numbers, the two numbers which we use in our calibration are the distance of the entrance pupil from the front principal plane denoted as  $H_1 E_n(a_n)$  and the distance of the exit pupil from the rear/back principal plane denoted as  $H_2 E_x(a_x)$ . Simple computations from Fig. 10 lead to  $a_n = 6.5$  mm and

Table 1. Calibration results on real data. (TL = Thin-lens, PC = Pupil-centric, ME = Moving Entrance, NF = Non-frontal, F = Frontal)

Method		Heikkila [7]	GenneryA [5]	GenneryB [5]	Kumar [10]	Ours
distortion model		Radial+Decentering	Radial	Radial	Radial	—
imaging model		TL	TL	TL+ME	PC	PC+ME
sensor orientation		F	NF	NF	NF	NF
#intrinsic parameters		(8)	(8)	(10)	(8)	(8)
scale	$\frac{s_y}{s_x}$	1.000	1.001	1.001	0.999	0.999
	$\lambda_p$ (mm)	8.240	8.383	8.358	8.650	8.593
principal point([7])/	$I_0$	218.647	223.612	223.859	229.662	228.645
CoD([5, 10])(pixels)	$J_0$	330.477	327.462	327.300	332.216	332.409
radial	$k_1$	−0.0019	−1.9e − 03	−4.3e − 03	−0.002	—
distortion	$k_2$	.000034	4.2e − 05	3.8e − 05	.00004	—
entrance pupil	$\epsilon_1$	—	—	8.122	—	−5.304
movement	$\epsilon_2$	—	—	15.178	—	6.474
decentering	$p_1$	.000015	—	—	—	—
distortion	$p_2$	−.000085	—	—	—	—
image sensor	$\alpha$	—	0.105	0.114	−0.424	−0.451
rotation (degrees)	$\beta$	—	0.424	0.439	3.278	3.051
re-projection error		0.077	0.079	0.078	<b>0.076</b>	<b>0.076</b>

Table 2. Std. deviation of calibration parameters shown in Table. 1.

Calibration Method	$\lambda_{px} = \frac{\lambda_p}{s_x}$	$\lambda_{py} = \frac{\lambda_p}{s_y}$	$I_0$	$J_0$	$\alpha$	$\beta$
Heikkila [7]	0.955	0.944	0.457	0.387	—	—
GenneryA [5]	<b>0.131</b>	<b>0.127</b>	0.559	0.625	0.029	0.027
GenneryB [5]	0.173	0.170	0.411	0.431	<b>0.022</b>	<b>0.023</b>
Kumar [10]	0.674	0.650	1.042	0.336	0.038	0.230
Ours	1.176	1.180	<b>0.288</b>	<b>0.239</b>	0.031	0.083

$a_x = 31.4$  mm. We use these values in our calibration method in Eq. 6 and Eq. 7.

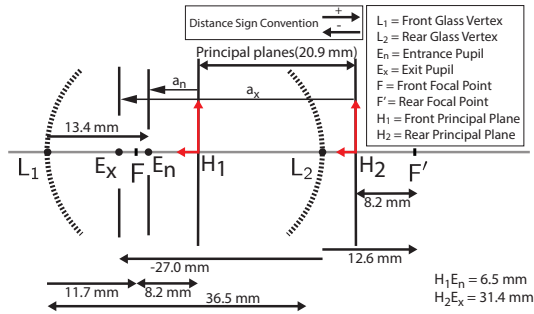


Figure 10. Lens data-sheet values for Cinegon 1.4/8 mm Lens.

**Analysis of calibration results:** Here, we present the results of proposed calibration method and compare the pixel re-projection error with the representative algorithms for each of the calibration categories mentioned in Sec. 4. The results for these prior techniques and our current method (last column) are shown in Table. 1. The different imaging conditions have been abbreviated in the caption of Table. 1.

The second column of this table corresponds to the implementation of Heikkila [7]. Here, the estimate of the CoD (218.647, 330.477) corresponds to the principal point on

the image plane where the effective optic axis is normal to the sensor. In the third and fourth column, we shown the results obtained from a Category 2 calibration method of Gennery [5]. Here, we implement two variations of their method. The method in column labeled as GenneryA uses a thin-lens imaging model with radial distortion and non-frontal sensor and a fixed entrance pupil location. The fixed entrance pupil allows us to conduct calibration over the same number of intrinsic calibration parameters, namely 8, as in our proposed method. The re-projection error here is higher than all other methods. The estimate of sensor tilt of  $\approx 0.424$  degrees is also far from the known lens specifications. The method in GenneryB is a full implementation of [5] where we have thin-lens imaging, moving entrance pupil, non-frontal sensor and radial distortion parameterization. The entrance pupil model adds two more calibration parameters making the number of intrinsic parameters 10. Thus, the re-projection error of GenneryB is less than GenneryA since we have used more number of parameters, yet it is more than our method. In the fifth column, we implement the calibration method of Kumar [10] from Category 3, where pupil-centric imaging model is used along with non-frontal sensor and radial distortion. Since decentering is encompassed in non-frontal sensor model, it is not calculated. The re-projection error is 0.076 pixels. Finally,

the sixth column presents the calibration results from our method where there is no explicit radial or decentering distortion modeling. It can be seen that among the calibration methods with same number of intrinsic parameters, our method achieves re-projection error of 0.076 pixels which is equal (approximated to three decimal places) to the current state of the art [10]. By obtaining re-projection error comparable to the state-of-the-art methods, we confirm the validity of our proposed model. The standard deviation of a set of intrinsic calibration parameters is also shown in Table. 2. The deviation in the estimation of CoD is minimal in our case. An interesting observation is that the standard deviation of sensor tilt angle  $\beta$  for Kumar [10] is 0.230 which is close to the difference of  $\beta$  estimates of 3.278 degree and 3.051 degree obtained by them and our current method. Thus, our estimates of sensor tilt ( $\alpha, \beta$ ) have better confidence levels.

## 7. Image Undistortion

In our paper, an ideal undistorted image is the one which is formed when all the image rays from the scene pass through the same entrance pupil location and thus have a fixed magnification as given by Eq. 3. This entails predicting the intersection of red dotted line, corresponding to the image ray from  $P_w$  which passes through the ideal entrance pupil location  $E_n$  in Fig. 11, with the image sensor plane. But, the depth of  $P_w$  is along the actual distorted image ray (solid blue line in Fig. 11), is not known. Thus, the location of ideal undistorted point becomes depth dependent (solid green and blue lines in Fig. 11) and can not be predicted accurately from our calibration model. But, empirically we have observed that the variation in the position of undistorted image point (corresponding to solid blue line in Fig. 11) as a function of its depth is very small. Thus, we propose to use a fixed scene depth for all image points in the scene and then obtain undistorted image points. Fig. 11(middle) compares the straight line fitting error for image undistortion using our method for various assumed scene depths and the traditional polynomial method of [7]. The ground truth depth of checkerboard is in the range of 70 – 100 mm from the entrance pupil for which our undistortion error is 0.028 pixels (see the minimum in the red curve) which is smaller than obtained by traditional methods which is 0.030 pixels (blue line). For other depth ranges, the fitting error from our method is higher but upper bounded by 0.035 pixels. Fig. 11(bottom) shows a pair of distorted and corresponding undistorted image using our method and known approximate depth of 70 – 100 mm.

## 8. Summary

In summary, Hecht [6] has shown that for thin lenses, the location of aperture stop dictates the amount by which an ideal single image ray distorts. But in a real image, due to thick lens, we observe that different image rays get

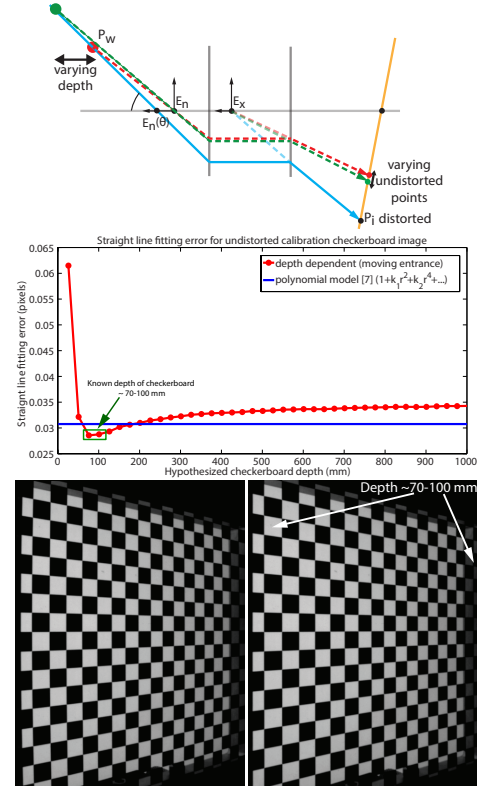


Figure 11. (top)Depth dependent undistortion (best viewed in color). (middle)Straight line fitting error using proposed (assuming different depths) vs traditional [7] undistortion. (bottom) Distorted image (left) is undistorted (right) using proposed method and assuming a depth of 75 mm.

distorted by different amounts. This would imply multiple aperture stop locations dictating the distortion of each image ray. This, we know is physically impossible as aperture stop is fixed in an imaging system. But, now if we assume that thick lenses follow pupil-centric image formation, then the location of entrance pupil, which is the image of aperture stop, dictates the amount of distortion. And, it has been shown in fish-eye literature [5], that the entrance pupil location can vary as a function of incidence image ray angle and is not fixed. This allows us to optically explain the occurrence of varying amounts of radial distortion due to varying entrance pupil location. We then propose a new calibration model considers pupil-centric imaging with moving entrance pupil and a non-frontal sensor. The moving entrance pupil models radial distortion and non-frontal sensor models decentering distortion.

## 9. Acknowledgments

This support of the US Office of Naval Research (ONR) grant N00014-12-1-0259 is gratefully acknowledged.



## References

- [1] M. Aggarwal and N. Ahuja. A pupil-centric model of image formation. *IJCV*, 2002. 3, 5
- [2] J.-Y. Bouguet. Camera calibration toolbox for matlab. Website, 2000. [http://www.vision.caltech.edu/bouguetj/calib\\_doc/](http://www.vision.caltech.edu/bouguetj/calib_doc/). 6
- [3] D. Brown. Decentering distortion of lenses. *Photogrammetric Engineering*, 32(3):444–462, May 1966. 1, 2, 5
- [4] A. E. Conrady. Decentered lens-systems. *Monthly Notices of the Royal Astronomical Society*, 79:384–390, 1919. 1
- [5] D. Gennery. Generalized camera calibration including fish-eye lenses. *IJCV*, 2006. 1, 2, 3, 4, 5, 6, 7, 8
- [6] E. Hecht. *Optics*. Addison-Wesley, 4th edition, 1998. 2, 3, 5, 6, 8
- [7] J. Heikkilä and O. Silven. A four-step camera calibration procedure with implicit image correction. In *CVPR*, 1997. 1, 2, 4, 5, 7, 8
- [8] S. O. Inc. Cinegon 1.4/8mm compact. Website, 2000. <https://www.schneideroptics.com>. 6
- [9] A. Kumar and N. Ahuja. A generative focus measure with application to omnifocus imaging. In *ICCP*, 2013. 6
- [10] A. Kumar and N. Ahuja. Generalized pupil-centric imaging and analytical calibration for a non-frontal camera. In *CVPR*, 2014. 1, 3, 4, 5, 6, 7, 8
- [11] T. Melen and J. G. Balchen. Modeling and calibration of video cameras. In *SPIE*, 1994. 6
- [12] R. A. Newcombe and A. Davison. Live dense reconstruction with a single moving camera. In *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*, pages 1498–1505, June 2010. 6
- [13] R. Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *IJRA*, 1987. 1, 4, 5
- [14] J. Weng, P. Cohen, and M. Herniou. Camera calibration with distortion models and accuracy evaluation. *PAMI*, 1992. 1, 2, 4, 5
- [15] N. Xu and et al. Generating omnifocus images using graph cuts and a new focus measure. In *ICPR*, volume 4, pages 697–700. IEEE, 2004. 6
- [16] Z. Zhang. A flexible new technique for camera calibration. *PAMI*, 2000. 1, 2, 4, 5