Non-frontal Camera Calibration using Focal Stack Imagery

Avinash Kumar, Narendra Ahuja
Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign
Email: {avinash,n-ahuja}@illinois.edu

***Abstract*—A non-frontal camera has its lens and sensor plane misaligned either due to manufacturing limitations or an intentional tilting as in tilt-shift cameras. Under ideal perspective imaging, a geometric calibration of tilt is impossible as tilt parameters are correlated with the principal point location parameter. In other words, there are infinite combinations of principal point and sensor tilt parameters such that the perspective imaging equations are satisfied equally well. Previously, the non-frontal calibration problem (including sensor tilt estimation) has been solved by introducing constraints to align the principal point with the center of radial distortion. In this paper, we propose an additional constraint which incorporates image blur/defocus present in non-frontal camera images into the calibration framework. Specifically, it has earlier been shown that a non-frontal camera rotating about its center of projection captures images with varying focus. This stack of images is referred to as a focal stack. Given a focal stack of a known checkerboard (CB) pattern captured from a non-frontal camera, we combine geometric reprojection error and image bur error computed from current estimate of sensor tilt as the calibration optimization criteria. We show that the combined technique outperforms geometry-only methods while also additionally yielding blur kernel estimates at CB corners.**

I. Introduction and Related Work

Camera calibration is the task of estimating the intrinsic and extrinsic parameters of a camera imaging a 3D scene and capturing 2D images of this scene on the image sensor. The intrinsic parameters encode the physical characteristics of the camera and the extrinsic parameters determine the 3D pose of the camera with respect to a known world coordinate system. In this paper, we focus on calibrating non-frontal cameras whose sensor and lens plane are not constrained to lie on the same plane. Typically, all cameras can be considered as non-frontal given that we take manufacturing limitations into account. Although, sometimes an intentionally tilted sensor can be used for tilt-shift photography, focal stack acquisition [1] and omnifocus imaging [2].

Previously, a number of techniques have been proposed for calibration of such cameras including that of [3]–[7] which do not take sensor tilt into account and [8]–[10] which also estimate the sensor tilt. As these techniques are inherently geometric in nature as then minimize the pixel-reprojection error to obtain calibration estimates, they will be referred to as *geometric* techniques henceforth.

Additionally, other image properties e.g. image blur, vignetting have also been used for camera calibration. In [11], a flat texture surface was imaged and the effect of vignetting in the captured image was used for calibration. In [12], geometric and blur properties of a lens were used to model the point spread function (PSF) of the lens and then the model parameters were optimized from the observed PSF images. In [13], blurred edges in a CB image were used to estimate the radius of a circular blur kernel as well as the location of the CB corner. These measurements were then compared with physically modeled predictions parameterized by the calibration parameters, and the resulting error minimized to obtain the optimal values of the parameters.

This paper combines ideas from geometric and image blur based methods to achieve non-frontal camera calibration. While [13] focuses on the problem of detecting corners under unwanted blurring of the pinhole CB images, we propose that instead of treating blur as unwanted, a sequence of intentionally blurred images, in addition to traditional, sharp images, can provide useful constraints to handle some inherent ambiguities in calibrating cameras modeled as being non-frontal. These ambiguities pertain to the highly correlated parameters of sensor-tilt angle and principal point location and have been discussed and analyzed in [9], [10] who also provide geometric solutions. In this work, we propose an additional blurring constraint to solve this ambiguity problem. We leverage on the idea that the sensor tilted about the optic axis of the camera produces unique image blur pattern on the captured image of a scene from a non-frontal camera. Thus, if the image blur can be analyzed to uniquely estimate the sensor tilt and the principal point.

Now, geometric calibration requires sharp pinhole images while the blur constraint requires presence of image blur in the input calibration image data. While both constraints cannot be satisfied at the same time for an image, we propose to use a focal stack as an input. A focal stack has each scene point imaged with varying amounts of focus including zero blur in a sequence of images. Also, it has been shown in [1] that a non-frontal camera rotating about its optic center can be used to capture a focal stack. Thus, we have blurred images from a non-frontal camera. An omnifocus imaging technique [2] can be applied to the focal stack to compute a sharp focused image. This image is a representative of an ideal pinhole image and is used as an input to geometric calibration framework.

Thus, our proposed calibration technique takes a focal stack as input, warps and registers them with respect to each other such that a CB corner across the focal stack appears at the same pixel location in a global image coordinate frame. The registered images are then used to compute an omnifocus image [2] from which sharp CB corners are detected [14]. Given the registration parameters, the omnifocus image is then

warped back to each of the focal stack images to synthesize a focused focal stack which can now be used in a conventional geometric calibration framework [9] to obtain extrinsic and intrinsic parameter estimates. While the extrinsic parameter encodes the object distance of a particular CB corner, the intrinsic parameters of sensor tilt encodes the image distance of that particular CB corner. Thus, using the thin lens equation and a Gaussian blur kernel model, the parameters of the blur kernel can be computed and applied to the focused focal stack to synthesize a conventional focal stack. The synthetic conventional focal stack can then be compared with the observed conventional focal stack to give the image blur error. The combined geometric and image blur error is then used to optimize all calibration parameters including blur kernel parameters.

We describe the geometry of blurred image formation in Sec. II-A. The focal stack acquisition technique is described in Sec. II-B. Sec. II-C describes pinhole image formation in a non-frontal camera and the calibration parameters. Sec. II-D explains the computation of blur kernel from calibration parameters. The proposed calibration approach is described in Sec. III. Finally, Sec. IV shows the results and the efficacy of the proposed calibration technique on real data sets.

II. Image Formation, Modeling and Acquisition

A. Defocus Image Formation

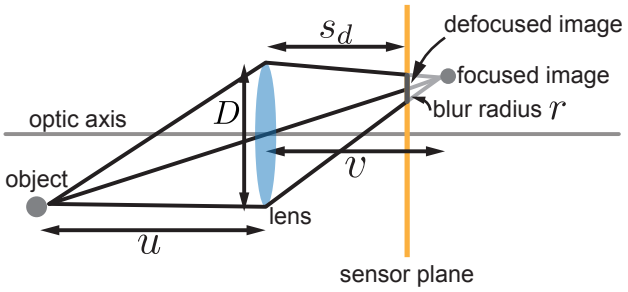


Fig. 1. Defocus image formation.

Fig. 1 shows image formation with aperture wide open in a thin-lens setting where an object is imaged on an image sensor. From the thin lens law [15], we have that the distance u of the object and the distance v of its sharply focused image from the thin lens are conjugate to each other. This implies that they are related by the thin-lens equation

$$\frac{1}{u} + \frac{1}{v} = \frac{1}{F}, \tag{1}$$

where F is the focal length of the lens. In the defocused camera setting, when the image sensor is moved to a distance s_d from the lens which is different than v, the captured image of the object is blurred. If the object is a point source of light, then this image is commonly referred to as the point spread function (PSF). The shape of the PSF for spherical lenses is typically assumed to be circular with radius parameter r and light intensity distribution $h(x, y)$ is assumed to be Gaussian [16]. A functional form of $h(x, y)$ can be obtained as follows:

$$r = \frac{D}{2}\left|\frac{s_d}{v} - 1\right| \qquad \text{and} \qquad h(x,y) = \frac{1}{2\pi\sigma^2}\exp -\frac{x^2+y^2}{2\sigma^2}. \tag{2}$$

where D is aperture diameter of the lens and is assumed to be known a priori based on F-number [15], $\sigma \approx \frac{r}{c}$ where c is camera dependent (empirically set to 0.3 in this paper).

B. Focal Stack Images as Non-frontal Camera Calibration Data

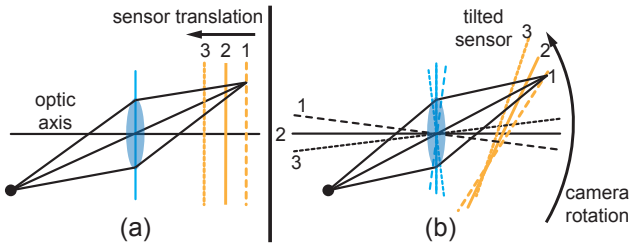


Fig. 2. Focal stack acquisition by (a) moving sensor along the optic axis. (b) rotating a camera with tilted sensor about the optic center [1]. On comparison of the imaging geometry in both cases we observe that the images formed on sensor location 1,2 and 3 are similar. Thus, both systems can be used interchangeably to generate similar focal stack images. In our work we use the configuration (b).

A focal stack is a collection of focused and defocus images of a static scene captured by varying some camera setting before each image capture and then acquiring the image. These settings could be e.g. relative distance between lens and sensor, object distances from the lens or aperture size. While calibrating a camera, it is usually assumed that the camera configuration is fixed. Thus the ideal way to capture a focal stack for non-frontal camera calibration is to translate the CB along an axis parallel to the optic axis of the lens [17]. The conjugate technique of keeping the CB fixed and moving the sensor along the optic axis (See Fig. 2(a)) will also generate a similar focal stack. We will refer to it as the *conventional* focal stack acquisition method. In [1], a further abstraction of the conventional technique was proposed where the sensor of the camera was first tilted with respect to the lens plane. Then, this camera was rotated about the optic center. This resulted in a single scene point getting imaged with varying amounts of focus in each captured image (See Fig. 2(b)). This happened because while rotation of the camera caused the object distance to change in each image, the sensor tilt at each rotation caused the image distance to change as well. Thus, by the thin-lens Eq. 1, both u and v varied to generate defocus, while in the conventional case only image distance v varied to generate defocus. Although tilted sensor camera [1] gives more freedom for focal stack acquisition, it has the same defocusing properties as conventional technique. Thus, without any loss of generality, the camera proposed in [1] can be used to acquire focal stack calibration data.

C. Geometric Imaging Model

Now, we describe the geometric image formation image projection equations as a function of various calibration parameters for a non-frontal camera [1]. Compared to conventional cameras with frontal sensor, the only added calibration parameter for non-frontal cameras are two Euler angle rotation parameters of the image sensor [8]. For calibration, we consider the following four coordinate systems (See Fig. 3):

1) World coordinate system (C_W) located on one of the corners of the CB pattern,
2) Lens coordinate system (C_L) centered at the pinhole projection with its z axis aligned with the optic axis and the xy plane parallel to the lens plane.
3) Sensor coordinate system (C_S) located on the image sensor with origin at the location where the optic axis intersects with the sensor plane
4) Image coordinate system (C_I) in which the observed image points are defined.

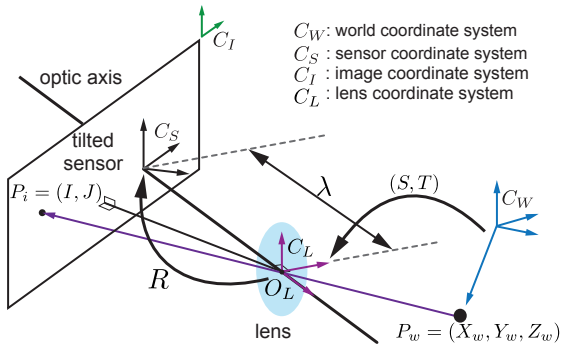


Fig. 3. Geometric image formation.

Assuming noiseless and distortion less imaging, we consider an object point $P_w = (X_w, Y_w, Z_w)$ in C_W and its corresponding measured image point $P_m = (I, J)$ in C_I. These two points can be mapped in terms of various camera calibration parameters as follows.

If C_W and C_L are related by a three Euler angle parameter rotation matrix $S_{=}(s_{ij} : 1 \leq (i,j) \leq 3)$ and translation $T = (t_x, t_y, t_z)$, then P_W can be expressed in C_L as $P_{C_L} = (x_l, y_l, z_l)$:

$$\begin{bmatrix} x_l \\ y_l \\ z_l \\ 1 \end{bmatrix} = \begin{bmatrix} S & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} P_W \\ 1 \end{bmatrix},$$

$$= \begin{bmatrix} s_{11}X + s_{12}Y + s_{13}Z + t_x \\ s_{21}X + s_{22}Y + s_{23}Z + t_y \\ s_{31}X + s_{32}Y + s_{33}Z + t_z \\ 1 \end{bmatrix}, \quad (3)$$

Let C_S be translated by $T_\lambda = (0, 0, \lambda)$ along the optical axis from the origin of C_L and rotated by rotation matrix R parameterized by two Euler angles [8]. P_L can then be expressed in the coordinate system C_S as

$$\begin{bmatrix} x_s \\ y_s \\ z_s \\ 1 \end{bmatrix} = \begin{bmatrix} R & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} I & T_\lambda \\ 0 & 1 \end{bmatrix} \begin{bmatrix} P_L \\ 1 \end{bmatrix},$$

$$= \begin{bmatrix} r_{11}X + r_{12}Y + r_{13}Z + r_{13}\lambda \\ r_{21}X + r_{22}Y + r_{23}Z + r_{23}\lambda \\ r_{31}X + r_{32}Y + r_{33}Z + r_{33}\lambda \\ 1 \end{bmatrix}, \quad (4)$$

where $R = (r_{ij} : 1 \leq (i,j) \leq 3)$. Since R is parameterized by two Euler angles, we note that $r_{21} = 0$ in Eq. 4. Given P_S and ignoring distortion, the intersection of the light ray, connecting P_S and the origin of C_L, with the non-frontal sensor plane can be obtained using simple coordinate geometry. This is the point on the sensor where the ideal image of the scene point is formed. Let this point be denoted as $P_{nf} = (x_{nf}, y_{nf})$. It can then be transformed to obtain final predicted image coordinates $P_p = (I_p, J_p)$ as:

$$I_p = \frac{x_{nf}}{s_x} - I_0; \quad J_p = \frac{y_{nf}}{s_y} - J_0 \quad (5)$$

where (I_0, J_0) are the location of origin of C_S in image coordinates and (s_x, s_y) are the size of the pixel in distance metric. Under ideal noiseless and distortion less imaging $P_m = P_p$. Finally, the set of 12 calibration parameters which need to be estimated are:

$$U = \Big[\underbrace{S}_{3\ \texttt{Euler ang.}}, \underbrace{T}_{3}, \underbrace{R}_{2\ \texttt{Euler ang.}}, \lambda, s_x, I_0, J_0 \Big] \quad (6)$$

where we assume that s_y is given to us as s_y decides the reference scale with respect to which calibration parameters are estimated.

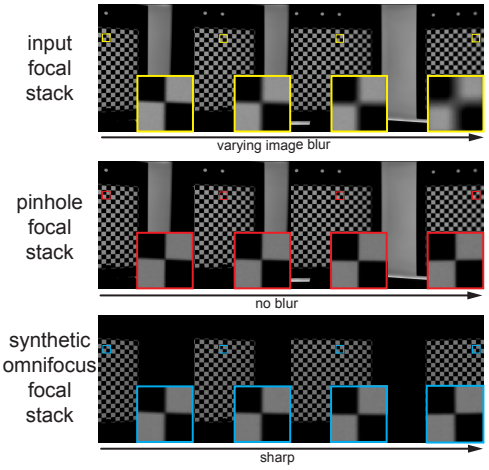


Fig. 4. Input: (top row) Input focal stack images (middle row) Corresponding pinhole focal stack used for computing registration parameters to warp input focal stack(See Sec. III-A). Output: (bottom row) Synthetic omnifocus focal stack obtained after omnifocus imaging (Sec. III-B) and dewarping (Sec. III-C). The camera was rotated from left to right. The inset shows details of intensity distribution at one single image location across image sequences in each row.

D. Integrating geometric and blur cues for calibration

Now, from the viewpoint of camera calibration, it can be seen that the blur parameters in Eq. 2 are dependent on camera calibration parameters of Sec. II-C as follows. Given the current estimate of calibration parameters (S, T) the checkerboard corners in lens coordinate system C_L are known from Eq. 3. Thus, the u parameter is determined for each CB corner. As optical focal length F is assumed to be given, v can be computed from lens equation Eq. 1. Also, given the current calibration estimate of R, λ, the distance of the a measured CB corner point P_m from the lens coordinate system C_L can

be computed. This amounts to computing s_d (Fig. 1). Given that D is already known before calibration and v, s_d has been computed, the blur radius r in Eq. 2 can be determined. This can be used to compute the current estimate of blur kernel $h(x, y)$ from Eq. 2 as a function of calibration parameters U and known parameters F, D.

III. Proposed Approach

The proposed approach consists of the following steps. Since, we are given only the focal stack images which are blurred, there is no guarantee that one of the images will have all the CB corners in focus. Thus, we need to combine all the focal stack images and obtain an omnifocus [2] image where all the CB corners are imaged in focus. The corners in this all focused image can then be reliably measured. But, in order to obtain an omnifocus image, the focal stack images need to be registered first in a global coordinate system which is addressed in Sec. III-A. The omnifocus image is then created in this global coordinate system as discussed in Sec. III-B. Since this image is created in a global coordinate system, it needs to be warped back to align with input focal stack images. This is discussed in Sec. III-C. Finally, for every input CB focal stack image, we obtain a corresponding sharp omnifocus image. We call this as *focused* focal stack. The two stacks are then used for combined geometric and blur cue based non-frontal camera calibration.

A. Registration of Focal Stack

The focal stack acquisition in Sec. II-B is dependent on the rotation of the non-frontal camera about a single viewpoint, which is the optic center O_L (Fig. 3). This has the benefit that consecutive focal stack images are related by a homography relation and can be registered. But, achieving a highly accurate unique viewpoint is a difficult task and there is always some parallax error between various camera poses due to which single viewpoint based homography doesn't exist and images are not registered accurately. The next best option is to compute planar homography between images of planar checkerboard pattern and then register images. But, due to presence of image distortion and blurring of features in consecutive focal stack images, the planar homography is not accurate enough to register images. Due to all these factors and the fact that even small inaccuracies in pairwise homography can accumulate into large errors while registering multiple images to a single image leads us to seek other techniques of image registration.

Since accurate registration is absolutely necessary for combining focal stack images and computing the omnifocus image, we treat registration as a preprocessing step done using conventional geometric camera calibration. Under this setting, the aperture of the non-frontal camera (Fig. 2(b)) is closed to maximum possible and then is rotated about O_L in same increments as were used to obtain focal stack images. This enables us to capture pinhole checkerboard images which are sharp and geometrically same as focal stack as shown in bottom row of Fig. 4(middle row). We call this *pinhole focal stack*. The pinhole focal stack is then used to do a conventional calibration using [14]. While, this calibration assumes that the sensor and the lens are parallel, the obtained parameters can be used obtain a numerical map M which projects all pinhole focal stack images to one global coordinate system where they are registered. The registration map M can then be used to register the actual focal stack images. The resulting registered focal stack images are shown in Fig. 5. Since the registered images have blurring due to interpolation, they are not used for any image processing. Rather, given a pixel location in the registered images, the map M is used to revert back to the original focal stack and that image is processed.

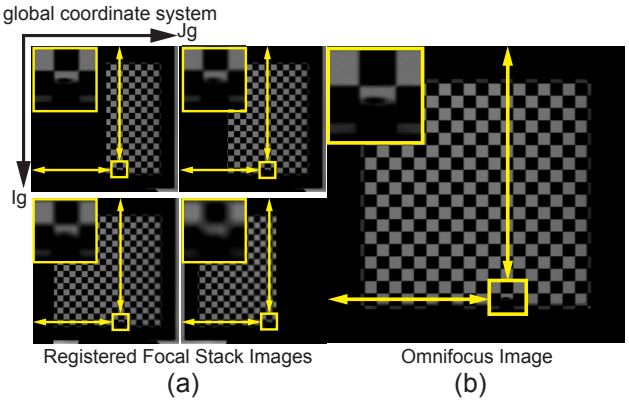


Fig. 5. (a) Input focal stack images of top row of Fig. 4 are registered to a global coordinate system (See Sec. III-A). A fixed pixel location is selected in each of the four images with inset showing the zoomed in intensity inside. Due to registration, they are aligned. (b) Omnifocus image obtained from the registered focal stack on the left (See Sec. III-B).

B. Computation of Omnifocus Image

Given the set of registered focal stack images, with each pixel being imaged in focus in at least one of the input images, an omnifocus image of the CB can be created [17]–[19]. Such an image denoted here as $I_{\texttt{omni}}$ has all the scene depths imaged in focus. See Fig. 5(b) for the omnifocus image obtained for the registered input data in Fig. 5(a). This image is obtained by computing a focus measure [2] at each pixel location across the focal stack. The focus measure computes the high frequency information in a window around a particular pixel location for all images in the focal stack. The image index which maximizes this measure is then selected as the image from which the focused pixel intensities are selected. Also, note that the focus measure is not computed on the registered images, rather on the actual focal stack images via the mapping M in Sec. III-A.

C. Dewarping and Corner Detection

Given the pixel wise mapping M between registered and input focal stack images (Sec. III-A), the omnifocus image $I_{\texttt{omni}}$ obtained in Sec. III-B, can be dewarped to input focal stack images with the only difference that the dewarped images are now focused. The dewarping technique requires the location of corresponding CB corner locations in $I_{\texttt{omni}}$ and the input focal stack images. The CB corners of focal stack are assumed to be same as that of the corresponding pinhole focal stack images (Sec. III-A). The CB corners in $I_{\texttt{omni}}$ are computed by a corner detection technique [14]. Now, since the input focal stack images have radial distortion and the registered images do not have any distortion, a linear mapping e.g. affine or homography is insufficient for accurate

dewarping. Thus, we apply a thin plate spline [20] technique to dewarp $I_{\texttt{omni}}$ into focused version of input focal stack images. These images are geometrically aligned to pinhole images shown in bottom row of Fig. 4, but differ in image blur. This difference arises from the fact that pinhole focal stack images had negligible Seidel aberrations (spherical aberration, coma, astigmatism) while these are inherently present in focused focal stack obtained by dewarping omnifocus image obtained from wide open aperture. Once the images are dewarped, the corners of the checkerboard are detected and stored. Thus, we have the geometric and blur information both obtained from focal stack images.

D. Combined Calibration Error Function

Given the current estimate of calibration parameters U (Eq. 6), the complete calibration error E_t is defined as a sum of geometric re-projection error E_g and the blur error E_p. Since images have distortion, the calibration parameter set U is appended with radial distortion terms. It was observed in [8] that if the sensor is modeled as non-frontal, then radial distortion about the optic center is sufficient to model all distortions. The radial distortion model consists of two calibration parameters: (k_1, k_2) relating a pair of distorted $P_S^d = (x_{sd}, y_{sd})$ and undistorted $P_S^u = (x_{su}, y_{su})$ image points in C_S as $P_S^d = P_S^u + g(P_S^u, \delta(k_1, k_2))$, where

$$g(P_S^u, \delta(k_1, k_2)) = \begin{bmatrix} x_{su}(k_1 r_{su}^2 + k_2 r_{su}^4) & y_{su}(k_1 r_{su}^2 + k_2 r_{su}^4) \end{bmatrix} \tag{7}$$

where $r_{su} = \sqrt{x_{su}^2 + y_{su}^2}$. Thus, the complete set of calibration parameters $U = (S, T, R, \lambda_{px}, \lambda_{py}, I_0, J_0, a_n, k_1, k_2)$, where $\lambda_{px} = \frac{\lambda_p}{s_x}$ and $\lambda_{py} = \frac{\lambda_p}{s_y}$.

E. Total Calibration Error

Let us assume that we have the estimate of calibration parameters as U. We also know the dimension and the location of corners on the CB in the world coordinate system C_W. The observations consists of two sets of data: (a) input focal stack which has CB images with varying amounts of blur, (b) a synthesized set of focused focal stack which has been generated by integrating (a) into an omnifocus image and dewarping the omnifocus image back to geometrically align with (a). The combined observation set is now used to design a combination of geometric and blur based error function which should minimize for optimal calibration parameters U^* as described below.

1) Geometric Error: Given the current estimate U of calibration parameters and the k^{th} known world point P_W^k, Eq. 5 gives the predicted image point $P_p^k = (I_p^k, J_p^k)$ on the image sensor. Additionally, we have measured image coordinates $P_m^k = (I_m^k, J_m^k)$ of CB corners from the focused focal stack images obtained after dewarping of omnifocus image (See Sec. III-C). The geometric error $E_g^k(U)$ for the k^{th} observed CB corner as a function of U can be defined as:

$$E_g^k(U) = (I_p^k(U) - I_m^k)^2 + (J_p^k(U) - J_k^k)^2. \tag{8}$$

2) Blurring Error: Given current estimate U, the blur kernel $h^k(x, y)$ at the k^{th} CB corner across all the focused focal stack images can be computed as explained in Sec. II-D. Also, we have the actual focal stack images (top row in Fig. 4) as input. Thus, at each k^{th} geometric corner location $P_m^k = (I_m^k, J_m^k)$ in each focused focal stack image, a square window of size $n \times n$ denoted as W_f around the corner can be blurred using $h^k(I_m, J_m)$ and compared with the corresponding observed $n \times n$ blurred corner W_b in the captured focal stack, then the blur error can be defined as:

$$E_p^k(U) = 1 - \texttt{NCC}(W_b^k, W_f^k * h^k(I_m, J_m)), \tag{9}$$

where NCC denotes the normalized cross correlating [21] between the two patches and its value lies between -1 and 1, where 1 denotes higher correlation. The use of NCC was justified as it is robust to intensity changes due to lens vignetting effects in the captured focal stack.

Thus, the total geometric and blur based error for all CB corner locations for all the images in the focal stack is

$$E_t(U) = \sum_{k=1}^{N} E_g^k(U) + E_p^k(U), \tag{10}$$

where N is the total number of CB corners in all the input focal stack images. This error is then minimized as:

$$U^* = \underset{U}{\operatorname{argmin}}\, E_t(U) \tag{11}$$

to get the optimal calibration parameters U^* using Levenberg-Marquardt algorithm [22].

IV. Experiments

A. Data Acquisition

The data is acquired using a custom made AVT Marlin camera, fitted with 1/2 inch Sony CCD sensor tilted by $\approx 3-4$ degrees with respect to the lens plane and a C-mount Schneider Cinegon 1.4/8mm compact lens. The acquired image resolution is 640×480. This camera is first centered empirically and the images of the checkerboard (CB) are captured by rotating the camera with a closed aperture (pinhole focal stack acquisition) and a wide aperture (focal stack imaging). In total, 5 images are captured in each setup without changing any other experimental conditions. A sample of 4 images is shown in Fig. 4(top row). The CB is custom made to get high positional accuracy of the corners. The size of each square in the CB is 5×5 mm.

B. Results: Real Data

TABLE I. Estimated calibration parameters and their standard deviation for 1 image.

Method		method (A)		method (B)	
		U^*	std. dev.	U^*	std. dev.
	$\lambda_x = \frac{\lambda}{s_x}$	847.812	4.007	846.806	0.259
	$\lambda_y = \frac{\lambda}{s_y}$	845.380	3.389	846.925	0.252
Principal	I_0	239.753	1.068	238.899	0.832
Point	J_0	330.773	1.065	332.159	0.679
Radial	k_1	-0.0023	$4.4e-05$	-0.0022	$3.9e-05$
	k_2	$5.7e-05$	$5.9e-06$	$5.6e-05$	$5.2e-06$
Tilt (deg)	$\cos^{-1}(r_{33})$	0.750	6.658	4.260	0.336
Re-projection(Geom. Err.) E_g		0.044911	-	0.040575	-

We estimate U^* using two methods and show that the proposed geometric and blur based approach clearly outperforms geometric-only calibration by comparing $E_t(U^*)$ and the variance of U^*. The two methods are: (A) *Conventional:*

Geometric calibration using pinhole focal stack images where $E_t(U) = \sum_{k=1}^{N} E_g^k(U)$ and (B) *Proposed:* Geometric and Blur, where the CB corners are taken from focused focal stack images and blurred images are from input focal stack ($E_t(U)$).

The calibration estimates U^* and their corresponding standard deviation along with total geometric re-projection error, computed from methods (A) and (B) are shown in Table I and Table II for one and five images in the focal stack. From Table I, it is observed that the re-projection error $E_g(U^*)$ using proposed method (B) is much less than those obtained from method (A). Similarly, the standard deviation of the obtained estimates is least for the proposed technique. Also, comparing the estimates of tilt of the sensor, it is found that method (A) estimates the tilt to be 0.75 degrees which is very different from the specifications provided by the manufacturer. But the estimates from method (B) which is 4.260 degrees is much closer to the camera specifications. The tilt is a critical parameter in depth from focus/defocus techniques [1].

Next we analyze the results of using more images for calibration. In this case also, the best performer with respect to total error and standard deviation of estimated parameters is method (B). We also observe that for the proposed technique, only one image is sufficient to get accurate results. Finally, In Fig. 6, we plot the estimated blur circle r (Eq. 2(b)) obtained from the calibration results of Table. II using method (B), where the inset shows the blurred image and the predicted blur radius in detail. Since, the sensor is tilted, the left part of the image in Fig. 6(a) is focused while the right part is defocused. This behavior is clearly observed by the estimated blur kernel sizes at the CB locations in Fig. 6(a), where the size increase from left to right along the image horizontal.

TABLE II. ESTIMATED CALIBRATION PARAMETERS AND THEIR STANDARD DEVIATION FOR 5 IMAGES.

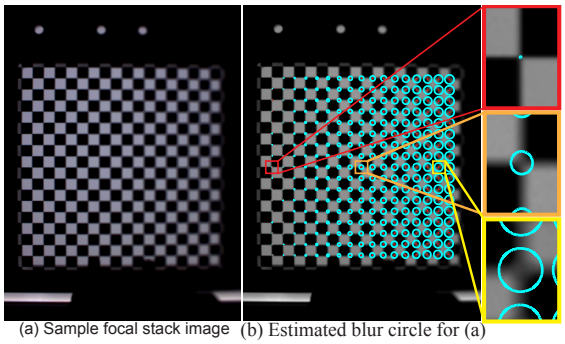
Method		method (A)		method (B)	
		U^*	std. dev.	U^*	std. dev.
	$\lambda_x = \frac{\lambda}{s_x}$	848.535	1.633	847.903	0.136
	$\lambda_y = \frac{\lambda}{s_y}$	847.062	1.579	847.984	0.129
Principal Point	I_0	238.804	0.531	239.338	0.383
	J_0	331.596	0.516	332.253	0.309
Radial	k_1	-0.0022	$3.5e-05$	-0.0022	$3.1e-05$
	k_2	$5.2e-05$	$4.3e-06$	$5.1e-05$	$3.7e-06$
Tilt (deg)	$\cos^{-1}(r_{33})$	2.487	2.856	4.232	0.166
Re-projection(Geom. Err.) E_g		0.045093	-	0.040671	-

V. CONCLUSION AND FUTURE WORK

In this paper, we have presented a framework for non-frontal camera calibration using geometric properties and image blur given a focal stack. We have shown improved results using the proposed technique in terms of the estimation error and the variance of the estimated parameters. The future work would be focused on making this technique independent of pinhole data by devising accurate image registration techniques to handle problems like parallax, radial distortion etc.

ACKNOWLEDGMENTS

This work was supported by US Office of Naval Research grant N00014-12-1-0259.



(a) Sample focal stack image (b) Estimated blur circle for (a)

Fig. 6. (a) Blurred focal stack image (b) Blur circle estimates from optimal U^* using Eq. 2 (inset shows details). The blurring of the CB corner looks consistent with the radius of the estimated Gaussian blur kernel. (best seen in color.)

REFERENCES

[1] A. Krishnan and N. Ahuja, "Range estimation from focus using a non-frontal imaging camera," *IJCV*, vol. 20, no. 3, pp. 169–185, 1996.

[2] A. Kumar and N. Ahuja, "A generative focus measure with application to omnifocus imaging," in *ICCP*, 2013, pp. 1–8.

[3] D. Brown, "Decentering distortion of lenses," *Photogrammetric Engineering*, vol. 32, no. 3, pp. 444–462, May 1966.

[4] R. Tsai, "A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses," *Robotics and Automation, IEEE Journal of*, vol. 3, no. 4, pp. 323 –344, August 1987.

[5] J. Weng, P. Cohen, and M. Herniou, "Camera calibration with distortion models and accuracy evaluation," *PAMI*, vol. 14, no. 10, pp. 965 –980, oct 1992.

[6] J. Heikkila and O. Silven, "A four-step camera calibration procedure with implicit image correction," in *CVPR*, jun 1997, pp. 1106 –1112.

[7] Z. Zhang, "A flexible new technique for camera calibration," *PAMI*, vol. 22, no. 11, pp. 1330 – 1334, nov 2000.

[8] D. B. Gennery, "Generalized camera calibration including fish-eye lenses," *IJCV*, vol. 68, pp. 239–266, 2006.

[9] A. Kumar and N. Ahuja, "Generalized pupil-centric imaging and analytical calibration for a non-frontal camera," in *CVPR*, 2014.

[10] A. Kumar and N. Ahuja, "Generalized radial alignment constraint for camera calibration," in *ICPR*, 2014.

[11] S. Kang and R. Weiss, "Can we calibrate a camera using an image of a flat,textureless lambertian surface?" in *ECCV*, 2000.

[12] Y. Shih, B. Guenter, and N. Joshi, "Image enhancement using calibrated lens simulations," ser. ECCV. Springer-Verlag, 2012, pp. 42–56.

[13] M. Baba, M. Mukunoki, and N. Asada, "A unified camera calibration using geometry and blur of feature points," in *ICPR*, 2006, pp. 816–819.

[14] J.-Y. Bouguet, "Camera calibration toolbox for matlab," Website, 2000, http://www.vision.caltech.edu/bouguetj/calib_doc/.

[15] E. Hecht, *Optics*, 4th ed. Addison-Wesley, 1998.

[16] B. Horn, *Robot Vision*. McGraw-Hill, 1986.

[17] S. Nayar and Y. Nakagawa, "Shape from Focus," *PAMI*, vol. 16, no. 8, pp. 824–831, Aug 1994.

[18] M. Subbarao, T. Choi, and A. Nikzad, "Focusing techniques," *Journal of Optical Engineering*, pp. 2824–2836, 1993.

[19] N. Xu and et al., "Generating omnifocus images using graph cuts and a new focus measure," in *ICPR*, vol. 4. IEEE, 2004, pp. 697–700.

[20] F. L. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations," *PAMI*, vol. 11, no. 6, pp. 567–585, 1989.

[21] J. P. Lewis, "Fast normalized cross-correlation," 1995.

[22] D. P. Bertsekas, *Nonlinear Programming*. Belmont, MA: Athena Scientific, 1999.