Omnifocus image synthesis using lens swivel

Indranil Sinharoy, Prasanna Rangarajan, Marc P. Christensen
Department of Electrical Engineering, Southern Methodist University, 6251 Airline Road, Dallas, Texas 75275-0338
isinharoy@smu.edu, prangara@mail.smu.edu, mpc@lyle.smu.edu

Abstract: We present a simple technique for synthesizing an infinite depth of field image from a sequence of photographs captured while rotating a *symmetric* lens about the center of the entrance pupil. We discuss the feasibility conditions and provide a Zemax simulation that verifies the method.

OCIS codes: (110.1758) Computational imaging; (110.4155) Multiframe image processing; (110.0110) Imaging systems.

1. Introduction

Cameras have limited depth of field (DOF): when we photograph a scene, only a finite region of the object space appears sharp in the image. A large DOF is often desirable in applications such as optical microscopy, machine-vision, surveillance and general photography. Methods such as focus-stacking, wavefront coding, and plenoptic imaging solve this problem to various degrees. We present a new method to computationally create an infinite DOF image that involves blending a series of images captured while rotating a lens pivoted at the center of the entrance pupil.

We briefly review two concepts that provide background to our approach. In focus stacking [1], a form of epsilon photography, a stack of photographs is captured using a conventional camera (in which the lens, sensor and focused plane are parallel) while focusing at intermittent depths (Fig. 1 (a)). Although no single photograph has the complete scene in focus, the stack as a whole contains all parts of the scene in focus. An image with large DOF is synthesized by blending the in-focus regions from the photographs in the stack. Another technique, called Scheimpflug imaging [2], is commonly used to image tilted object surfaces. It involves tilting the lens and/or the sensor resulting in the plane of sharp focus (PoSF) in the object space to rotate. Thus, as shown in Fig. 1(b), the PoSF can be oriented along a particular direction of interest, which may extend to infinity within the field-of-view, resulting in infinite DOF along that direction. However, the DOFs along other directions are still finite. In this work, we combined these two ideas to devise a new class of epsilon photography in which the synthesized image has most of the object space in focus.

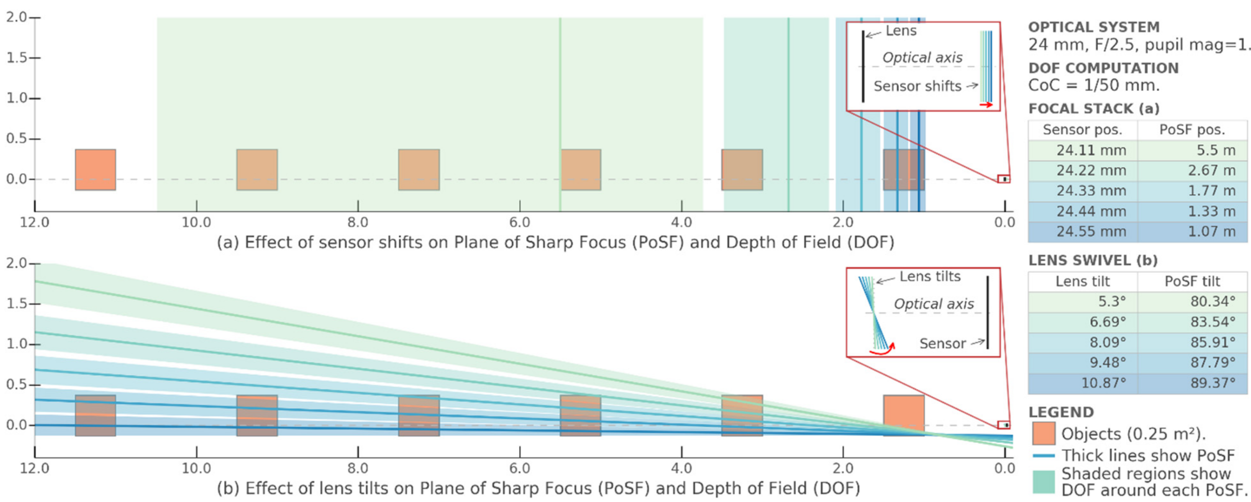

Fig. 1 PoSF and DOF: (a) Focus stacking using a conventional camera. (b) Scheimpflug imaging (for several tilts).

The main idea of this paper is that if we rotate a *symmetric* lens about the center of its entrance pupil, then the geometric warping in the images (mainly lateral translation and scaling) caused by the rotation of the lens is *independent* of the object coordinates. Concomitantly, rotating the lens forces the PoSF to swing through the three-dimensional object space (extending infinitely along the depth). Consequently, we can construct a stack with relatively few images that collectively contains most regions in focus within an infinitely extending depth as shown in Fig. 1(b). Since the image-image transformation between the images in the stack is independent of object coordinates, we can easily register and blend the photographs in the stack to synthesize an image exhibiting infinite DOF. Furthermore, if the camera is calibrated, we can use a closed form expression to register the images in the stack.

2. Theory

Fig. 2 shows a thick-lens model in which the origin of the coordinate frame $\{C\}$ pivots the lens. Similarly, the origin of image frame $\{I\}$ pivots the sensor. Object and image distances are measured from the entrance (E) and exit ($\acute{E}$) pupil centers respectively, which are themselves located at d_e and $\acute{d}_e$ from the lens' pivot point along the optical axis.

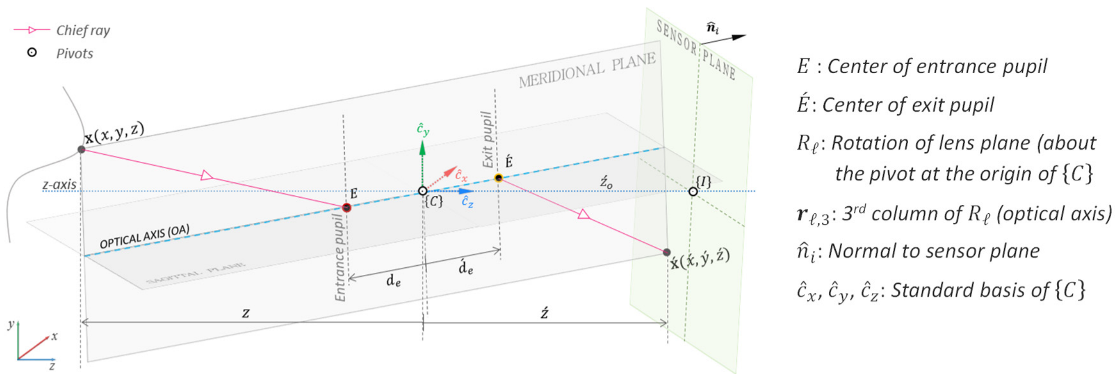

Fig. 2 Imaging model: Geometric image formation for arbitrarily oriented lens and sensor.

The geometric relation between a world point $\boldsymbol{x}$ and its corresponding image[1] $\acute{\boldsymbol{x}}$, derived in [3], is given as:

$$\acute{\boldsymbol{x}} = \acute{d}_e\, \boldsymbol{r}_{\ell,3} + \frac{\left(\hat{\boldsymbol{n}}_i(3)\acute{z}_o - \acute{d}_e \hat{\boldsymbol{n}}_i^T\, \boldsymbol{r}_{\ell,3}\right)}{\hat{\boldsymbol{n}}_i^T\, R_\ell M_p {R_\ell}^T \left(V\boldsymbol{x} + d_e\, \boldsymbol{r}_{\ell,3}\right)}\, R_\ell M_p {R_\ell}^T \left(V\boldsymbol{x} + d_e\, \boldsymbol{r}_{\ell,3}\right). \tag{1}$$

Where, R_ℓ is the rotation matrix applied to the lens; $\boldsymbol{r}_{\ell,3}$, the third column of R_ℓ, is a unit vector along the optical axis; $\hat{\boldsymbol{n}}_i$ is the sensor plane normal; $V = diag(1, 1, -1)$; $M_p = diag\left(1, 1, m_p\right)$, where, m_p is the pupil magnification. It is defined as the ratio of the exit pupil to the entrance pupil diameters. For symmetric lenses, $m_p = 1$. Symmetric lens, such as the Double Gauss and its variants, are commonly used in the design of Scheimpflug cameras.

Rotation of the lens about E results in a shift of the image-field along with a field dependent warp. For a given object point $\boldsymbol{x}$, the image points $\acute{\boldsymbol{x}}_m$ and $\acute{\boldsymbol{x}}_n$ observed under two instances of lens rotations are related as:

$$\acute{\boldsymbol{x}}_n = \left[p\, s\, R_{\ell;n} M_p {R_{\ell;n}}^T\, R_{\ell;m} M_p^{-1} {R_{\ell;m}}^T + \frac{d}{\hat{\boldsymbol{n}}_i(3)\acute{z}_o}\left(\boldsymbol{r}_{\ell,3;n} - p\, s\, R_{\ell;n} M_p {R_{\ell;n}}^T\, R_{\ell;m} M_p^{-1} {R_{\ell;m}}^T\right)\hat{\boldsymbol{n}}_i^T \right] \acute{\boldsymbol{x}}_m \tag{2}$$

Where, $p = \frac{(\hat{\boldsymbol{n}}_i(3)\acute{z}_o - d\, \hat{\boldsymbol{n}}_i^T \boldsymbol{r}_{\ell,3;n})}{(\hat{\boldsymbol{n}}_i(3)\acute{z}_o - d\, \hat{\boldsymbol{n}}_i^T \boldsymbol{r}_{\ell,3;m})}$, $s = \frac{(\hat{\boldsymbol{n}}_i^T R_{\ell;m} M_p {R_{\ell;m}}^T V \boldsymbol{x})}{(\hat{\boldsymbol{n}}_i^T R_{\ell;n} M_p {R_{\ell;n}}^T V \boldsymbol{x})}$, and d is the distance of the exit pupil from the entrance pupil. According to Eq. (2), the mapping between the two image points (for the same object point) depends on the object coordinates. However, if $m_p = 1$, then $s = 1$ for all R_ℓ and $\hat{n}_i$. Furthermore, if $\hat{\boldsymbol{n}}_i = [0,0,1]^T$; and if the lens rotates only about the x-axis by angle α; and if we designate the photograph acquired under no lens tilt as the reference image, then the mapping between the n^{th} instance and the reference ($m = 0$) is obtained from Eq. (2) as:

$$\acute{\boldsymbol{x}}_n = \begin{bmatrix} (d\cos\alpha - \acute{z}_o)/(d - \acute{z}_o) & 0 & 0 \\ 0 & (d\cos\alpha - \acute{z}_o)/(d - \acute{z}_o) & d\sin\alpha \\ 0 & 0 & 1 \end{bmatrix} \acute{\boldsymbol{x}}_0 \tag{3}$$

Eq. (3) suggests that the rotation of the lens causes a vertical shift by $(d\sin\alpha)$ and scaling by $\frac{(d\cos\alpha - \acute{z}_o)}{(d - \acute{z}_o)}$ of the image field, but most importantly, the transformation is independent of the object coordinates. This linear mapping allows us to register the images obtained under lens rotations before blending to generate an omnifocus image.

3. Simulation

Fig. 3(a) shows a schematic of the simulation we implemented in Zemax. A 24 mm, f/2.5 paraxial thick lens with $m_p = 1$ images three playing cards placed at 800mm, 1000mm and 1200mm from the lens' vertex. We used PyZDDE [4]

[1] In this model, the image point is defined as the point of intersection of the chief-ray from $\boldsymbol{x}$ with the image plane.

to automate the process of tilting the lens about the x-axis between $\pm 8°$ pivoted at the center of the entrance pupil and create a stack of 9 photographs. Fig. 3(b) is the photograph of the scene for $\alpha = -8°$. Observe that the individual images of the three cards were vertically shifted and de-magnified (not apparent in the figure) by the same amount, as predicted by Eq. (3). The in-focus regions, detected using a Laplacian of Gaussian (LoG) filter, are shown in Fig. 3(c). We registered the 13 photographs using the closed form expression Eq. (3) followed by blending the in-focus regions (as measured by LoG) from the photographs in the stack. Fig. 3(d) shows the synthesized image in which all three cards are in focus. Fig. 3(e) shows the degree of focus on the three planes in the composite image measured using the LoG filter. Amongst others, the LoG is commonly used as a focus measure filter in focus stacking.

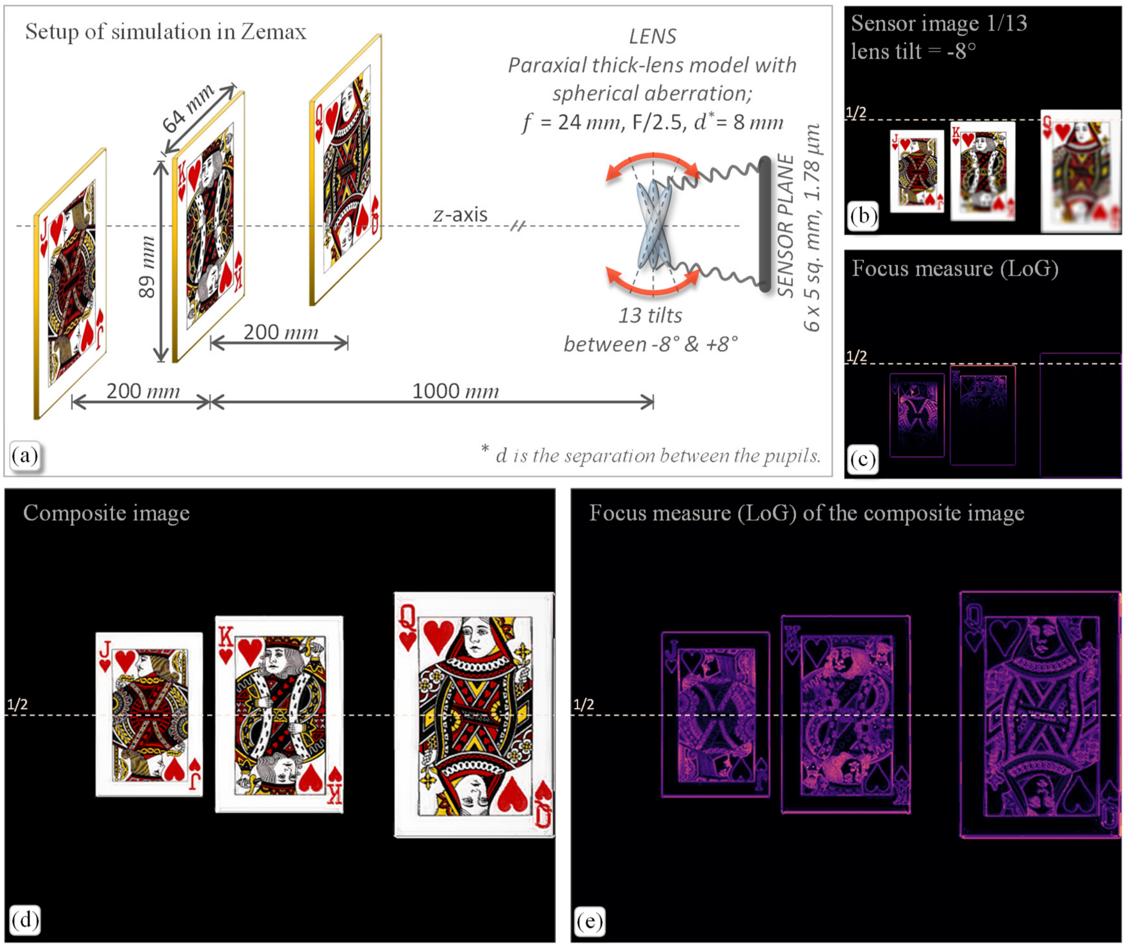

Fig. 3 Image simulation using Zemax and PyZDDE: (a) Setup. (b) Photograph for $\alpha = -8°$. (c) Focus-measure using LoG filter. (d) Resulting composite image. (e) Focus-measure of composite image.

4. Summary

We demonstrated a new type of epsilon photography for generating an omnifocus image by fusing a series of images captured under lens rotations. The crux of this simple method hinges on rotating a *symmetric* lens about the center of the entrance pupil. Also, if the camera is calibrated, the transformation for image registration is known in closed form.

5. References

[1] C. H. Anderson, J. R. Bergen, P. J. Burt, J. M. Ogden, "Pyramid Methods in Image Processing," RCA Engineer, vol. 29, pp. 33-41 (1984).

[2] Jacobson, Ralph, Sidney Ray, Geoffrey G. Attridge, and Norman Axford, *Manual of Photography* (Taylor & Francis, 2000), Chap. 10.

[3] Indranil Sinharoy, Prasanna Rangarajan, and Marc P. Christensen, "Geometric model of image formation in Scheimpflug cameras," PeerJ Preprints 4:e1887v1 https://doi.org/10.7287/peerj.preprints.1887v1 (2016).

[4] Indranil Sinharoy *et al.*, PyZDDE: Release version 2.0.2. Zenodo. 10.5281/zenodo.44295 (2016).