Image Destabilization: Programmable Defocus using Lens and Sensor Motion

Ankit Mohan[1] Douglas Lanman[2] Shinsaku Hiura[1,3] Ramesh Raskar[1]

[1]MIT Media Lab [2]Brown University [3]Osaka University

Abstract

We propose a novel camera setup in which both the lens and the sensor are perturbed during the exposure. We analyze the defocus effects produced by such a setup, and use it to demonstrate new methods for simulating a lens with a larger effective aperture size (i.e., shallower depth of field) and methods for achieving approximately depth-independent defocus blur size. We achieve exaggerated, programmable, and pleasing bokeh with relatively small aperture sizes such as those found on cell phone cameras. Destabilizing *the standard alignment of the sensor and lens allows us to introduce programmable defocus effects and achieve greater flexibility in the image capture process.*

1. Introduction

Nearly all modern digital cameras include an *image stabilization* feature to compensate for moderate camera shake during the exposure. A motion sensor (commonly a gyroscope) detects the direction of camera movement during the exposure, and the camera compensates for this by shifting either a lens element or the sensor in the opposite direction (commonly using piezoelectric actuators). While image stabilization enhances the sharpness of features within the intended plane of focus, it also degrades other parts of the scene. We exploit this property of image stabilization, and intentionally shift the sensor and lens in a synchronized manner to achieve useful optical effects. The related technique of panning is historically used to photograph moving objects. In this method the photographer rotates the camera during the exposure in order to track the motion of a moving object. If performed properly, the moving object appears stationary in the final image, and objects with different angular velocities than the camera body (e.g., a fixed background) appear motion-blurred.

Image stabilization, camera panning, and perspective control lenses attempt to compensate for the inherent limitations of a static sensor and fixed coaxial lens through various forms of actuation. The relative motion of three critical planes, (1) the scene, (2) the lens, and (3) the sensor, determines what parts of the scene appear in sharp focus on the sensor. Current commercial image stabilization systems only apply motion control to one of these planes, by either moving a lens element [9, 15] or the sensor [12, 18]. We propose a novel design in which both the lens and the sensor are perturbed during the exposure. In effect, by *destabilizing* the standard alignment of the sensor and lens, we can introduce programmable defocus effects and achieve greater flexibility in the image capture process which can potentially lead to new computational photography applications. In particular, we demonstrate new methods for simulating a lens with a larger effective aperture size (i.e., shallower depth of field) and methods for achieving approximately depth-independent defocus blur size by translating the lens and the sensor during the exposure.

(a) all-in-focus photograph with an $f/22$ lens

(b) *destabilized* photograph by shifting an $f/22$ lens and sensor

Figure 1. We destabilize the relative alignment of the pinhole and sensor during the exposure to achieve focus and depth of field effects using a pinhole camera. (a) Typical pinhole camera photo. (b) Image destabilization introduces a virtual focal length and a virtual aperture, effectively simulating a larger aperture lens.

Contributions: The primary contributions of this paper are as follows.

- We investigate and analyze the space of relative motions of a lens and sensor parallel to one another during an exposure. We show that such a setup can simulate the depth of field produced by a thin lens, with the corresponding *virtual focal length* and *virtual f-number* related to the ratio of the lens and sensor velocities.
- We propose a new *image destabilization* technique which uses these properties and builds on top of existing image stabilization hardware present in digital cameras. We analyze limitations and advantages of this technique.
- We propose a setup for achieving a shallower depth of field than possible with a given lens by simulating a lens with a larger virtual aperture size.
- We propose a setup where the captured photograph exhibits an approximately depth-independent defocus blur size within a user-specified depth range.

Limitations: Our prototype setup only allows for 1D translation of the lens and sensor, while 2D motion is necessary for a real system with a 2D lens aperture. As a result we only synthesize virtual 1D lens apertures in this paper. Several 2D motion configurations such as circular, elliptical, hypocycloidal, or spiral may be used to extend our techniques to achieve a virtual 2D lens. Although such trajectories only discretely sample the virtual aperture plane, we believe such sampling would be sufficient to approximately simulate 2D lenses for many applications.

1.1. Related Work

Exploiting coordinated movements of multiple imaging elements has a rich history in medical imaging. **Laminography** [4] is a technique typically applied in X-ray imaging to focus at distinct layers of a subject without using refractive elements. In contrast to more modern computed tomography methods, laminography directly forms a sharp, cross-sectional image using synchronized motion rather than post-capture computation. Laminography requires coordinated movements of any pair of the following three components: the radiation source, the subject, and the sensor. It is possible to focus on different layers of the subject by changing the relative velocities. The techniques proposed in this paper are conceptually similar to laminography – both exploit translation of two imaging elements in order to control the effective depth of field.

Coordinated translation of imaging elements is used when photographing moving scenes or when using perspective control lenses. Recently, Levin et al. [11] extended panning to photograph objects moving at different relative speeds. Instead of tracking a single motion, they constantly vary the angular velocity of the camera rotation during the exposure, resulting in a motion-invariant blur kernel, which can be deconvolved to recover a blur-free photo. Zomet and Nayar [20] used layers of programmable attenuating masks instead of a lens to simulate perspective control lenses by moving a pinhole to change the effective field of view. In a technique closely related to laminography, time delay and integration (TDI) [8] is used in CCD sensors to transfer stored charge between neighboring scanlines, producing a stable image of an object moving at a speed synchronized to the rate of charge transfer.

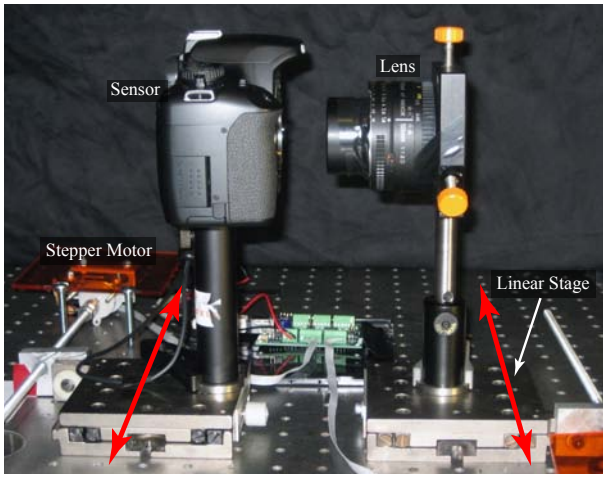

Figure 2. Our prototype setup for the image destabilization camera built from an off the shelf camera and lens mounted on motorized linear translation stages. The red arrows show the direction of motion of the sensor and lens.

A pleasing **bokeh** (appearance of defocused parts of the scene) is often more important than the sharpness and detail in the focused parts of the image, especially for portrait photography. Furthermore, shallow depth of field is often used for artistic and creative effects. The depth of field and the bokeh are closely-related to the shape and size of the aperture. Lenses with large apertures are as frequently desired for their shallow depth of field as for their light-gathering capability. Unfortunately, such lenses tend to be bulky and expensive. Small point-and-shoot cameras usually have more modest aperture sizes, greatly limiting their use for professional or creative photography. Computational methods, such as those introduced by Bae and Durand [1], can be used to hallucinate a shallower depth of field in such systems. In addition, Hasinoff and Kutulakos [5] demonstrated how to combine multiple images with varying aperture diameters to simulate a larger aperture. We propose a single exposure solution to synthesize narrower depth of field using the lens and sensor actuators already present for image stabilization.

Depth independent blur: In recent years, several authors have introduced methods for light-efficient capture of all-in-focus imagery. Wavefront coding [2] used a cubic phase mask to create an easily-invertible defocus kernel

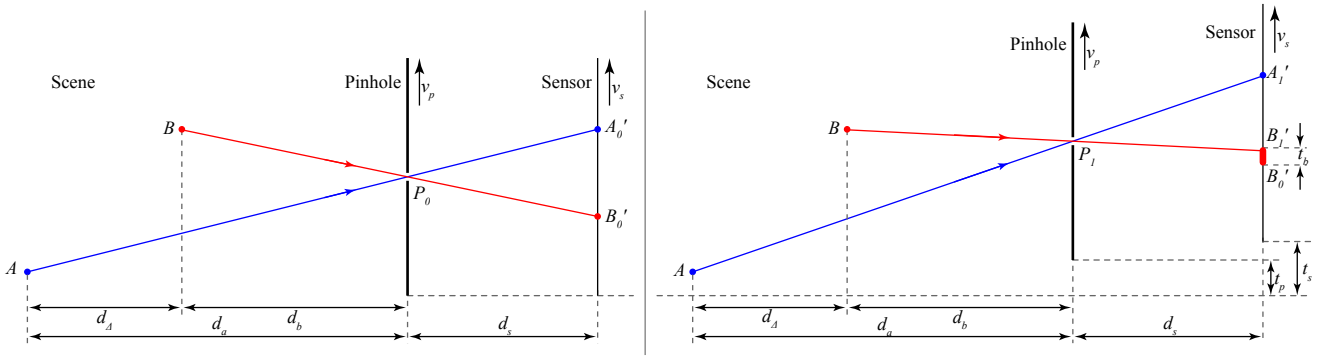

Figure 3. Ray diagram for a translating pinhole and sensor. (Left) Geometric configuration of a pinhole camera, where the pinhole at P_0 creates an image of the scene points A and B at points A'_0 and B'_0 on the sensor, respectively. (Right) Image formation for a translated pinhole at P_1, where the scene points A and B project to points A'_1 and B'_1 on the sensor, respectively. Using image destabilization the sensor displacement t_s is selected so that the images A'_0 and A'_1 coincide, producing a focused image of the plane containing A and blurring out all other scene planes, such as the plane containing B.

which is depth-invariant. In a similar design, Häusler [6] and Nagahara et al. [14] achieved a flexible depth of field by varying the distance between the lens and the scene or sensor, respectively, during the exposure. In these methods, the effective focal length is linearly-varied, either temporally or spatially, to ensure that each scene plane is focused at some point during capture. Conceptually this is similar to Levin et al. [11], where every angular velocity is synchronized with the camera rotation at some point during the exposure. We explore a similar solution to depth-invariant photography and demonstrate preliminary results for obtaining all-in-focus imagery using coordinated sensor and lens translation.

2. Defocus using a Shifting Pinhole and Sensor

Before considering coordinated motion of the lens and sensor, we briefly review the geometry of static pinhole photography. Figure 3(left) shows a simple ray diagram for a pinhole camera, where scene points A and B are projected to points A'_0 and B'_0 on the sensor, respectively. Since the pinhole selects a single ray from each scene point, the entire scene will appear focused on the sensor – irrespective of the distance of a given point from the pinhole. In this section we develop a novel technique, based on coordinated translations of the pinhole and the sensor, to obtain a virtual plane of focus and effective depth of field using a pinhole camera. While we derive the properties of a shifting pinhole and sensor in 1D, the theory in this section extends to 2D. Such extensions come at the price of more complex motion trajectories and possibly limited sampling of the virtual aperture plane.

2.1. Pinhole Camera Laminography

Similar to laminography, pinhole photography presents three components that can be shifted during the exposure: the scene, the pinhole, and the sensor. Unlike laminography, in photography we are interested in observing reflected light rather than transmitted radiation. As such, the pinhole effectively replaces the radiation source as the center of projection. We observe that concepts from laminography can be extended to pinhole photography in order to *reduce* the depth of field using synchronized translation of the pinhole and sensor.

As shown in Figure 3, we translate the pinhole parallel to the sensor with velocity v_p and translate the sensor with velocity v_s. We desire a constraint on these velocities such that the acquired image is *focused* on a specific scene plane at a distance d_a from the pinhole. We expect points in other scene planes will be *defocused*. Simple geometric constraints are adequate to determine the necessary pinhole and sensor velocities. As the pinhole moves from P_0 to P_1, the image of point A will shift from A'_0 to A'_1. To focus on the plane containing point A, the sensor must be translated such that A'_0 and A'_1 overlap. If t_p is the pinhole displacement and t_s is the sensor displacement, then $t_s = \left(1 + \frac{d_s}{d_a}\right) t_p$. Since this applies to any point on the plane at d_a from the pinhole, a coordinated parallel translation of the pinhole and sensor produces an image focused at d_a. Specifically, if the pinhole moves a constant velocity v_p during the exposure, then the sensor must translate with a constant velocity

$$v_s = \left(1 + \frac{d_s}{d_a}\right) v_p. \tag{1}$$

Now consider the scene point B at a distance d_b from the pinhole plane. The image of this point moves from B'_0 to B'_1 as the pinhole moves from P_0 to P_1. The total displacement t_b of the image of B as the pinhole translates over a distance t_p is given by

$$t_b = d_s \left| \frac{1}{d_b} - \frac{1}{d_a} \right| t_p.$$

In essence, *destabilizing* the relative alignment of the sensor and pinhole reduces the depth of field of our optical setup. For such a pinhole laminography configuration, the diameter of the circle of confusion is then given by

$$c_P = \frac{d_s}{d_a}\left(\frac{|d_\Delta|}{d_a + d_\Delta}\right) t_p, \tag{2}$$

where $d_\Delta = d_b - d_a$ is the distance from the plane of focus.

2.2. Comparison to a Thin Lens

We now compare the depth of field effects produced by pinhole laminography to that of an ideal thin lens. A thin lens is governed by the thin lens equation

$$\frac{1}{f_T} = \frac{1}{u} + \frac{1}{v}, \tag{3}$$

where f_T is the focal length of the thin lens, u is the object distance, and v is the image distance. Rearranging this expression and comparing with Equations 1 (where $d_a = u$, and $d_s = v$), the virtual focal length f_P for pinhole laminography is given by

$$f_P = \left(\frac{v_p}{v_s}\right) d_s. \tag{4}$$

The diameter of the circle of confusion [7] for a thin lens is given by the relation

$$c_T = \frac{f_T}{d_a - f_T}\left(\frac{|d_\Delta|}{d_a + d_\Delta}\right) A, \tag{5}$$

where A is the aperture diameter of the thin lens. Combining Equation 5 with the thin lens equation, we find

$$c_T = \frac{d_s}{d_a}\left(\frac{|d_\Delta|}{d_a + d_\Delta}\right) A.$$

Comparing this result with Equation 2, it is clear that the total displacement t_p in pinhole laminography must be equal to the aperture size A in order to replicate the circle of confusion for a given thin lens. Thus, the virtual f-number (the ratio of the virtual focal length to the virtual aperture size) for pinhole laminography is given by

$$N_P = \frac{f_P}{t_p} = \left(\frac{v_p}{v_s}\right)\left(\frac{d_s}{t_p}\right). \tag{6}$$

We conclude that synchronized translation of the pinhole and sensor allows a pinhole camera to replicate the effect of an arbitrary thin lens whose properties can be well modeled by geometric optics. Adjusting the relative translation velocities $\{v_p, v_s\}$ and total displacements $\{t_p, t_s\}$ of the pinhole and sensor allows the synthesis of a thin lens with focal length f_T and f-number N_T. This result can be better understood by interpreting a thin lens as a uniform array of translated pinholes and prisms, as discussed in Georgiev et al. [3]. Under this model, the image detected by the sensor is a linear superposition of the individual images formed by each shifted pinhole-prism pair. A local segment of the thin lens with focal length f_T, located a distance t_p from the optical axis, acts as a pinhole followed by a prism that produces a constant angular deflection $\alpha = -t_p/f_T$. Under the paraxial approximation, the prism effectively translates the resulting pinhole image by a distance t_s given by

$$t_s = -\alpha d_s = \frac{d_s t_p}{f_T} = \left(1 + \frac{d_s}{d_a}\right) t_p.$$

This translation is identical to the necessary sensor translation given by Equation 1. Pinhole laminography achieved by the synchronized translation of a pinhole and the sensor effectively creates a "thin lens in time", where the pinhole translation scans the aperture plane and the sensor translation replaces the action of the local prisms.

3. Defocus using a Shifting Lens and Sensor

While the defocus analysis in Section 2 is informative, a pinhole is rarely used in photography due to loss of light and diffraction. However, the theory can also be applied to a translating lens with a finite aperture size. A pinhole can be interpreted as a thin lens with an infinitely-small aperture located at the optical center. The virtual focal length and f-number for such a configuration is given by Equations 4 and 6 and the point spread function (PSF) is a box function for 1D motions (or a pillbox for 2D) corresponding to the circle of confusion in Equation 2. As we increase the aperture size, the overall PSF h^L is a combination of the virtual PSF due to pinhole and sensor translation and the physical PSF due to the lens aperture. The overall PSF is given by

$$h^L_{f_T,N_T,f_P,N_P}(d) = h^T_{f_T,N_T}(d) \star h^P_{f_P,N_P}(d), \tag{7}$$

where h^T is the physical PSF of the thin lens, h^P is the virtual PSF due to sensor and lens motion, and d is the distance of the point source from the lens plane. We conclude that translating a finite aperture lens synchronized with the sensor results in the creation of a second virtual lens, and the effective PSF of the resulting system is simply the convolution of the PSFs of the real and virtual lenses.

3.1. Defocus Enhancement

For the special case where the real and virtual focal lengths are matched (i.e., $f_T = f_P$), a shifting lens and sensor behaves very similar to a static lens of the same focal length, but with a larger effective aperture size (or smaller effective f-number). For this situation, a single plane is in focus and the size of the circle of confusion rapidly increases for scene points located away from this plane. The

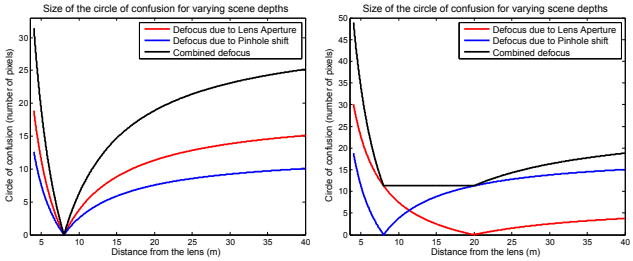

Figure 4. (Left) Defocus enhancement using a lens with focal length $50mm$, aperture $15mm$, focused at $8m$, and total lens displacement of $10mm$. The cumulative blur size is approximately equal to the sum of the two for any scene depth. (Right) Depth-invariant blur size with the lens focused at $20m$, and a $15mm$ total lens displacement. The cumulative blur size is approximately a constant for all distances in the range of $8m$ to $20m$.

increased effective aperture size yields a depth of field that is shallower than what is obtained by either a static lens with f-number N_T or a pinhole laminography configuration with f-number N_P. The overall f-number N_L of a shifting lens and sensor is given by

$$\frac{1}{N_L} = \frac{1}{N_T} + \frac{1}{N_P},$$

where N_P is the virtual f-number given by Equation 6. Even though the effective aperture size is increased, the total light entering the camera during the exposure remains identical to that allowed by the unmodified physical aperture.

The effective PSF of a shifting lens and sensor is the convolution of the real and virtual PSFs. As a result, limitations of the physical PSF due to the lens can be addressed by engineering an appropriate virtual PSF using laminography. The laminographic component $h^P(d)$ depends on the relative velocities and paths (in 2D) of the lens and sensor as they translate. These parameters may be easier to control than the optical elements within the lens. Since coordinated translation introduces additional blur, laminography can be applied to attenuate high-frequency components in the physical PSF and improve the overall bokeh.

Figure 4(left) shows a plot of the size of the circle of confusion for a fixed lens in red, pinhole laminography in blue, and the combined effect by shifting a lens and sensor in black. The overall size of the combined circle of confusion is approximately equal to the sum of the two cases, and the depth of field is shallower for the combination.

3.2. Depth-Invariant Blur Size

The more general case where $f_T \neq f_P$ results in a setup that cannot be duplicated with only a single fixed lens and sensor. In this case the two focusing mechanisms do not focus at identical planes. As a result, no single plane is focused on the sensor and the effective PSF for any scene

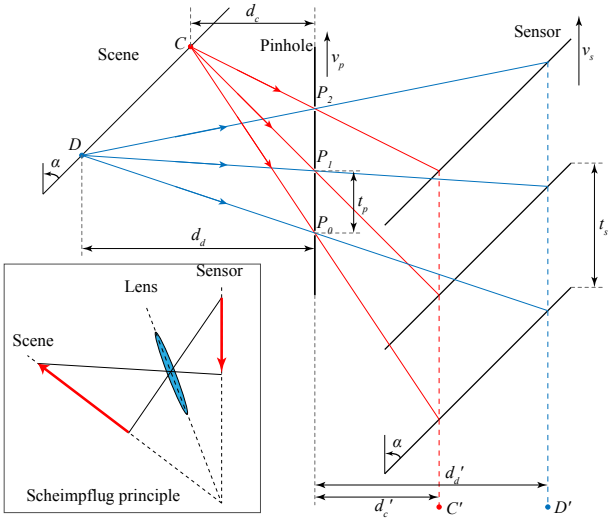

Figure 5. Ray diagram for tilted sensor. The pinhole moves with a velocity v_p, and the tilted sensor moves parallel to it with a velocity v_s. In-focus points C and D lie on a plane that is parallel to the sensor. The overall effect achieved is similar in spirit to the traditional Scheimpflug setup for a tilted lens and sensor (inset).

depth is the convolution of the two individual PSFs for that depth. If d_a^T and d_a^P are the two in-focus planes for a physical lens and laminography, respectively, then the size of the combined circle of confusion is approximately constant for all planes that lie between them, as shown in Figure 4(right). This results in a depth-invariant blur size for the specified range of scene distances. This is similar to the effect attained by wavefront coding [2] and Nagahara et al. [14], and an all-in-focus image may be obtained by deconvolving the constant blur kernel. Unfortunately, even though the size of the circle of confusion is invariant to the depth, the PSF varies with depth in general. However, for the special case of a Gaussian PSF for both the real and virtual focusing mechanisms, laminography may be used to obtain an overall depth-invariant Gaussian PSF for the combined setup.

4. Tilted Sensor

So far we have considered the case where the sensor is parallel to the lens or the motion direction of the pinhole. The case where the sensor and lens are static and not parallel to one another is well understood by the Scheimpflug principle [13, 10]. The plane of focus for such a setup is not parallel to either the lens or the sensor, and passes through the line of intersection formed by the extended planes containing the lens and the sensor as shown in Figure 5(inset).

Unfortunately, the Scheimpflug principle cannot be reproduced by shifting the pinhole/lens and the sensor as we did in the previous sections. This is because the virtual focal length for the pinhole laminography configuration, as shown in Equation 4, is a function of the pinhole-sensor

separation d_s. While this does not affect the case where the sensor is parallel to the direction of the pinhole motion, the virtual focal length varies over the surface of a tilted sensor, thus violating the traditional Scheimpflug principle. However, it is possible to obtain similar results using the configuration shown in Figure 5. The sensor is tilted at an angle α and moves *parallel* to the direction of the pinhole. Two points C and D focus on the image sensor over time, and we wish to find the geometric relationship between these points. From the figure, we obtain

$$\frac{d_c}{d_c + d'_c} = \frac{d_d}{d_d + d'_d} = \frac{t_p}{t_s} = \frac{v_p}{v_s}.$$

This gives the relation $\frac{d_c}{d'_c} = \frac{d_d}{d'_d}$, which implies that the line joining in-focus points C and D is parallel to the sensor (due to similar triangles). The setup focuses on a plane that is parallel to the sensor. The exact plane of focus depends on the ratio of the sensor velocity to the pinhole velocity $\frac{v_s}{v_p}$, and we can use Equation 1 to find it.

Similar to Section 3, we can use a translating lens in place of a translating pinhole. A lens parallel to the sensor also focuses on a plane parallel to the sensor (the exact plane depends on the focal length of the lens). Once again, we can either match the virtual focal length to the physical focal length to enhance the bokeh, or keep it distinct to produce a depth-invariant blur size across a tilted scene.

5. Implementation and Results

Simulating the defocus effects of a thin lens requires two dimensional displacement of the lens and sensor. We describe an initial design using one-dimensional displacements, observing that the theory equally applies to this case and that our results confirm the overall concepts developed in this paper.

Prototype: We constructed the image destabilization prototype shown in Figure 2, containing a 12.2 megapixel Canon EOS Digital Rebel XSi camera, a Nikkor $50mm$ $f/1.8D$ lens with manual aperture control, and a pair of linear translation stages driven by external stepper motors. We attached a second diverging lens behind the Nikkor lens to form a focused image on the sensor. The camera and lens were enclosed in a box to prevent stray light from reaching the sensor. The stepper motors and camera were computer controlled and we ensured that the exposure occurred outside of the stepper motor ramp-up and ramp-down phases. The translation stages supported a total displacement of $4cm$ and typical exposures ranged from 5 to 30 seconds.

Pinhole-based Destabilization: We demonstrate image destabilization using a pinhole aperture by stopping-down the lens to $f/22$. As shown in Figures 1 and 6, we are able to change the virtual focal length by adjusting the velocity ratio as per Equation 4, allowing various scene planes to be brought into focus. Adjusting the total displacement according to Equation 6, we are also able to control the virtual f-number of the virtual lens. Figure 7 shows results for a scene containing an array of point sources at varying depths – the size of the circle of confusion varies with depth as expected.

Figure 6. Reduced depth of field using image destabilization and a physical aperture of $f/22$. (Top) Focused on the front figure with $t_p = 10mm$, (Middle) focused in the middle with $t_p = 5mm$, and (Bottom) focused in the back with $t_p = 10mm$. The relative translation velocities are adjusted according to Equation 1 to focus at various planes in the scene. The f-number reduces with increasing lens translation range, t_p.

Lens-based Destabilization: As shown in Figure 8, we reduce the depth of field using the method outlined in Section 3. Our prototype seems to improve the aesthetic quality of the bokeh since the coordinated translation effectively applies a low-pass filter that removes high-frequency artifacts due to spherical aberration.

(a) static lens with an $f/22$ aperture

(b) image destabilization with a virtual focal plane in the center

(c) static lens with an $f/2.8$ aperture, focused in the center

(d) shallower depth of field by translating an $f/2.8$ lens and sensor

(e) simulating an aspheric lens with a vertical slit aperture on a translating lens and sensor

Figure 7. PSFs observed in mirror ball reflections. The spheres are placed at increasing distances from the lens (from left to the right), and are illuminated by a single bright point light source. The PSF due to image destabilization is 1D because our translation is restricted to 1D. In (e), the PSF changes from horizontal for points close to the camera (due to the virtual aperture), to vertical for points farther away (due to the physical lens).

We achieve a depth-invariant blur size by matching the physical blur kernel due to the lens aperture, and the virtual blur kernel due to translating lens and sensor. We place a horizontal slit on the lens to make the PSF purely one dimensional. The lens is physically focused on the closest figure from the camera, and the virtual focal plane is at the farthest figure. As shown in Figure 9, the resulting blur size is approximately depth-invariant, allowing the application of non-blind image deconvolution to recover an all-in-focus image.

The physical and virtual blurs also can be made orthogonal to produce strongly depth-dependent PSF [16]. We cover the entrance aperture of the lens with a vertical slit. Such a configuration produces the strong astigmatism as shown in Figure 7. In this result, points close to the camera (left side of the photo) exhibit horizontal blur due to the virtual aperture, whereas points farther away (right side of the photo) show vertical blur due to the physical aperture. The advantage of our method over using an aspheric lens is that we use the same lens to also obtain an ordinary photograph by simply changing the $\frac{v_s}{v_p}$ ratio.

Figure 8. Reducing the depth of field using image destabilization. (Top) A photograph taken with a static lens with an $f/2.8$ aperture (Bottom) Synchronized translation of the lens and sensor simulates the effect of a larger virtual aperture. The depth of field is shallower, and the bokeh is visually more pleasing both in front of and behind the plane of focus.

6. Discussion and Future Work

Limitations: Our analysis is based on translation of a single thin lens, while most cameras use compound lenses with a single moving element for image stabilization. Our prototype moves the lens and sensor at relatively slow speeds (about $1 - 2mm/s$), and higher speeds would be needed for practical applications. Also, our analysis ignores any scene motion during the exposure and results would be similar to those obtained when using image stabilization in a dynamic scene. Finally, digital sensors are designed to work best for light arriving approximately orthogonal to the sensor plane. By shifting the sensor relative to the lens we break this assumption, possibly introducing additional vignetting for large offsets.

Future Work: Due to limitations of the prototype, only parallel movements of the lens and sensor have been demonstrated. It is an outstanding problem to analyze the more general case of arbitrary directions of movement and lens orientations. In our experiments we move the lens and sensor at constant velocities; modulating the velocity profile over time should allow the overall PSF of the system to be shaped and to allow control of the bokeh characteristics. There is also a possibility to achieve non-planar focal surfaces using non-linear motion of the sensor and the

Figure 9. Depth-invariant blur size. (Top) We create a depth invariant blur size by focusing a lens with a horizontal slit aperture close to the camera, and create a virtual focal plane farther in the scene using image destabilization. The PSF produced by both forms of defocus is very similar and the overall result is an approximately depth invariant defocus kernel. (Bottom) Although the overall PSF is not strictly invertible, Richardson-Lucy deconvolution provides reasonable results for an approximate all-in-focus image.

lens. Though we obtain a depth-independent blur size, the resulting PSF is not easily invertible. Shaping the PSF, by combining the ideas of fluttered shutter [17] and coded aperture [19], might make deconvolving the cumulative defocus blur a well-posed problem. Such systems that combine the effects of motion blur, lens defocus, and computational focusing are promising avenues of future research.

Conclusion: In this paper we use motion blur (by intentionally shifting the sensor and lens) to produce effects similar to lens defocus. We combine two different types of defocus (motion-based and lens-based) to produce interesting optical effects that may be hard to achieve otherwise. While similar results might be achieved with light field cameras [3, 19], capturing the complete 4D data is unnecessary, and to some degree excessive, for the applications we present. Furthermore, our setup captures high-resolution photos and does not suffer from discrete angular sampling artifacts. While our hardware setup might seem somewhat unusual, it can be easily implemented using existing image stabilization technology already present in most cameras. We believe that greater control over the sensor and lens position during the exposure period offers enhanced flexibility useful for computational photography applications.

Acknowledgements

We thank the reviewers and Gabriel Taubin for useful suggestions, Grace Woo for video preparation, and Quinn Smithwick for assistance in constructing the prototype.

References

[1] S. Bae and F. Durand. Defocus magnification. In *Eurographics*, volume 26, pages 571–579, 2007. 2

[2] W. T. Cathey and E. R. Dowski. New paradigm for imaging systems. *Applied Optics*, 41(29):6080–6092, 2002. 2, 5

[3] T. Georgiev, K. C. Zheng, B. Curless, D. Salesin, S. Nayar, and C. Intwala. Spatio-angular resolution tradeoffs in integral photography. In *EGSR*, pages 263–272, 2006. 4, 8

[4] S. Gondrom, S. Schröpfer, and D. Saarbrücken. Digital computed laminography and tomosynthesis. In *Computerized Tomography for Industrial Applications and Image Processing in Radiology*, 1999. 2

[5] S. W. Hasinoff and K. N. Kutulakos. A layer-based restoration framework for variable-aperture photography. In *IEEE ICCV*, pages 1–8, 2007. 2

[6] G. Häusler. A method to increase the depth of focus by two step image processing. *Optics Comm.*, 6(1):38–42, 1972. 3

[7] E. Hecht. *Optics*. Addison Wesley, 2001. 4

[8] G. C. Holst. *CCD Arrays, Cameras and Displays*. JCD Publishers; SPIE Optical Engineering, 1998. 2

[9] E. Kawakami, Y. Niwa, Y. Ogino, M. Yoshii, S. Suda, and M. Ohwada. Anti-vibration imaging device. US Patent 4780739, 1988. 1

[10] A. Krishnan and N. Ahuja. Range estimation from focus using a non-frontal imaging camera. *IJCV*, 20(3):169–185, 1996. 5

[11] A. Levin, P. Sand, T. S. Cho, F. Durand, and W. T. Freeman. Motion-invariant photography. In *SIGGRAPH*, 2008. 2, 3

[12] S. Masuda and Y. Hara. Position detector, camera-shake compensation mechanism, and image capture apparatus. US Patent 7132824, 2006. 1

[13] H. M. Merklinger. *Focusing the View Camera*. Seaboard Printing Limited, 1996. 5

[14] H. Nagahara, S. Kuthirummal, C. Zhou, and S. K. Nayar. Flexible depth of field photography. In *ECCV*, 2008. 3, 5

[15] K. Oizumi, N. Kitagishi, and S. Yamzaki. Optical system for stabilizing an image. US Patent 5270857, 1993. 1

[16] R. Piestun, Y. Y. Schechner, and J. Shamir. Propagation-invariant wave fields with finite energy. *JOSA A*, 17(2):294–303, 2000. 7

[17] R. Raskar, A. Agrawal, and J. Tumblin. Coded exposure photography: motion deblurring using fluttered shutter. In *SIGGRAPH*, pages 795–804, 2006. 8

[18] S. Seo. Stage apparatus and camera shake correction apparatus using the same. US Patent 7319815, 2008. 1

[19] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin. Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing. In *SIGGRAPH*, volume 26. ACM, 2007. 8

[20] A. Zomet and S. K. Nayar. Lensless imaging with a controllable aperture. In *IEEE CVPR*, pages 339–346, 2006. 2