An Analysis of Focus Sweep for Improved 2D Motion Invariance

Yosuke Bando
TOSHIBA Corporation
yosuke1.bando@toshiba.co.jp

Abstract

Recent research on computational cameras has shown that it is possible to make motion blur nearly invariant to object speed and 2D (i.e., in-plane) motion direction, with a method called focus sweep that moves the plane of focus through a range of scene depth during exposure. Nevertheless, the focus sweep point-spread function (PSF) slightly changes its shape for different object speeds, deviating from perfect 2D motion invariance. In this paper we perform a time-varying light field analysis of the focus sweep PSF to derive a uniform frequency power assignment for varying motions, leading to a finding that perfect 2D motion invariance is possible in theory, in the limit of infinite exposure time, by designing a custom lens bokeh used for focus sweep. With simulation experiments, we verify that the use of the custom lens bokeh improves motion invariance also in practice, and show that it produces better worst-case performance than the conventional focus sweep method.

1. Introduction

Defocus and motion deblurring has been an active area of research in the computational camera community, where a computational deblurring process is facilitated by a camera-hardware-assisted smart capture process. One of the important categories of such capture processes is *invariant capture*, which makes the point-spread function (PSF) invariant to scene depth or motion, thereby bypassing the need for PSF identification [11, 9, 13, 8, 7, 6]. Recently, it has been shown that the *focus sweep* method, which sweeps the plane of focus through a range of scene depth by moving the lens or the image sensor along the optical axis during exposure, not only makes defocus blur invariant to object depth [9, 13], but also makes motion blur nearly invariant to object speed and 2D (i.e., in-plane) motion direction [3, 17]. While focus sweep is shown to be near-optimal both in terms of invariance and high-frequency preservation for certain combinations of depth and motion ranges, it is still suboptimal. In particular, the focus sweep PSF slightly changes its shape for different object speeds even in theory.

In this paper we perform an analysis of focus sweep capture to explore the possibility of achieving perfect 2D motion invariance at least in theory, and of achieving better performance in practice. Building upon the time-varying light field analysis for joint defocus and motion blur PSFs in [3], we find a way to uniformly distribute PSF frequency power over all possible scene depths and motions within some predetermined ranges. This uniform assignment leads to perfect 2D motion invariance in theory, in the limit of infinite exposure time, in a similar manner to the case of an accelerating camera [11] that achieves perfect 1D (e.g., horizontal) motion invariance with the infinite exposure assumption. We also show that this uniform assignment is achieved by a custom lens bokeh used for focus sweep. Deblurring simulation verifies that the use of the custom lens bokeh improves motion invariance in practice and produces better worst-case performance than using the standard lens.

The presented analysis inherits the assumptions and limitations of the previous work [3]. Namely, we assume that scenes are Lambertian; scene depth and motion have limited ranges; and object motions are in-plane (no z-axis motion) and constant (no acceleration) within the exposure time.

2. Related Work

We only summarize previous invariant capture methods and analyses here. For depth-invariant capture, several methods have been proposed ranging from the use of a cubic phase plate [8] to focus sweep [9, 13], to an annular diffuser at the aperture [7], and to a chromatically-aberrated lens [6]. They are all designed in a way that light rays impinging on a sensor pixel are uniformly distributed over a range of depth, so that every depth is fractionally focused. An analysis of depth invariance of computational cameras can be found in [2].

For motion-invariant capture, it is shown that blur can be made invariant to object speed in an *a priori* chosen 1D (say, horizontal) direction by translating the image sensor horizontally with a constant acceleration [11]. Comparisons with a (non-invariant) motion deblurring method [15] can also be found in [1].

Researchers have found that focus sweep also provides

near motion invariance in addition to depth invariance, and that the near motion invariance holds for arbitrary 2D (in-plane) motion directions [3, 17]. However, the motion invariance is only approximate even in theory.

3. Analysis

In this section we first briefly summarize the time-varying light field analysis of the conventional focus sweep in [3] in Sec. 3.1, and we derive a uniform frequency power assignment over depth and motion ranges in order to theoretically achieve perfect 2D motion invariance in Sec. 3.2, which we show can be realized as modified focus sweep with a custom lens bokeh in Sec. 3.3.

3.1. Preliminaries

We model a degraded image $D(\mathbf{x})$ as 5D convolution of an incoming time-varying light field $l(\mathbf{x}, \mathbf{u}, t)$ with a kernel $k(\mathbf{x}, \mathbf{u}, t)$ as

$$D(\mathbf{x}_0) = \iiint k(\mathbf{x}_0 - \mathbf{x}, -\mathbf{u}, -t) \cdot l(\mathbf{x}, \mathbf{u}, t) d\mathbf{x} d\mathbf{u} dt, \quad (1)$$

where we use two-plane parameterization for the light field [12] with $\mathbf{x} = (x, y)$ denoting locations on the image sensor and $\mathbf{u} = (u, v)$ locations on the aperture, and t denotes time (see Fig. 1). The integrals are taken over $(-\infty, +\infty)$. The kernel k represents how each light ray passing through the point $\mathbf{u}$ on the aperture at time t is mapped to the position $\mathbf{x}$ on the sensor, and for the conventional focus sweep capture,

$$k(\mathbf{x}, \mathbf{u}, t) = \delta(\mathbf{x} - wt\mathbf{u}) R(|\mathbf{u}|/A) R(t/T), \quad (2)$$

where R is a rect function such that $R(z) = 1$ for $|z| < 1/2$ and $R(z) = 0$ otherwise. $R(t/T)$ indicates that the shutter is open during exposure time T for $t \in [-T/2, T/2]$, $R(|\mathbf{u}|/A)$ indicates that the aperture is open inside the disc with diameter A, and $\delta(\mathbf{x} - wt\mathbf{u})$ indicates that a ray passing through $\mathbf{u}$ on the aperture is mapped to $\mathbf{x} = wt\mathbf{u}$ on the sensor with the focused depth wt changing along time at *focus sweep speed* w.

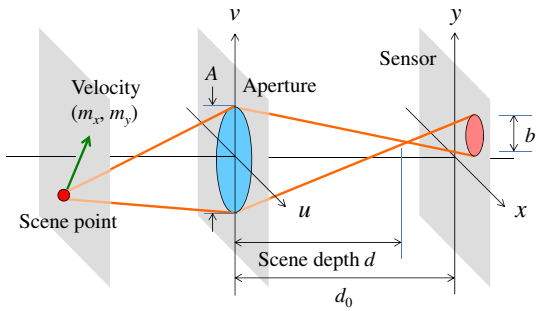


Figure 1. Light field parameterization $xyuv$ and a moving scene point. Scene depth d is taken as a distance from the aperture towards the sensor. Reproduced from [3].

Now, let d and $\mathbf{m} = (m_x, m_y)$ be depth and motion (velocity) of a scene point. The PSF representing joint defocus and motion blur can be written as

$$\phi_{s,\mathbf{m}}(\mathbf{x}) = \iint k(\mathbf{x} + s\mathbf{u} + \mathbf{m}t, \mathbf{u}, t) d\mathbf{u} dt, \quad (3)$$

where $s = (d - d_0)/d$, with d_0 denoting the distance between the aperture and the sensor, encodes the object depth in a way that it corresponds to the slope in the 4D light field space [5, 14, 11]. The optical transfer function (OTF) of this PSF is given as

$$\hat{\phi}_{s,\mathbf{m}}(\mathbf{f_x}) = \hat{k}(\mathbf{f_x}, -s\mathbf{f_x}, -\mathbf{m} \cdot \mathbf{f_x}), \quad (4)$$

where Fourier transform is denoted using a hat symbol and $\mathbf{f_x} = (f_x, f_y)$ represents frequency in the x and y directions. This means that the OTF is a 2D slice of the 5D Fourier transform $\hat{k}(\mathbf{f_x}, \mathbf{f_u}, f_t)$ of the kernel k with the following assignments

$$\mathbf{f_u} = -s\mathbf{f_x}, \qquad f_t = -\mathbf{m} \cdot \mathbf{f_x}, \quad (5)$$

where $\mathbf{f_u} = (f_u, f_v)$ and f_t represent frequency in the u, v and t directions.

The squared modulation transfer function (MTF, the magnitude of OTF, which characterizes deblurring performance [11]) of the conventional focus sweep PSF is derived as follows by explicitly taking 5D Fourier transform of Eq. (2).

$$|\hat{\phi}_{s,\mathbf{m}}(\mathbf{f_x})|^2 = \begin{cases} \dfrac{A^2}{w^2|\mathbf{f_x}|^2}\left(1 - \dfrac{4|\mathbf{m} \cdot \mathbf{f_x}|^2}{A^2 w^2 |\mathbf{f_x}|^2}\right) \\ \quad (\text{for } |s| \le \frac{Tw}{2}, \ |\mathbf{m} \cdot \mathbf{f_x}| \le \frac{Aw}{2}|\mathbf{f_x}|) \\ 0 \qquad (\text{otherwise}) \end{cases} \quad (6)$$

Hence, within some scene depth range S and motion range M such that $|s| \le S/2$ $(\le Tw/2)$ and $|\mathbf{m}| \le M/2$ $(\le Aw/2)$, the MTF of the conventional focus sweep does not depend on scene depth s, but it gradually falls off for faster motion $|\mathbf{m}|$ (see the red plot in Fig. 2). While this fall-off can be minimized by setting the motion range as $M = Aw/\sqrt{3}$, the motion invariance is still approximate.

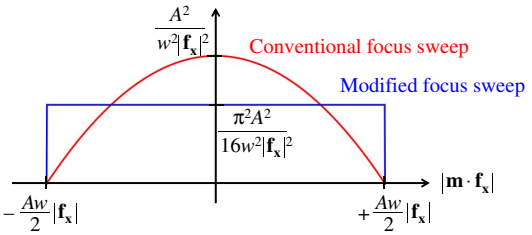


Figure 2. Squared MTF of the conventional focus sweep PSF (red, Eq. (6)) and that of the modified focus sweep PSF (blue, Eq. (13)), plotted with respect to the axis corresponding to object speed.

3.2. Derivation of an Improved Focus Sweep Kernel

We begin by examining the derivation of Eq. (6). We first take the 5D Fourier transform of the conventional focus sweep kernel Eq. (2) as

$$\hat{k}(\mathbf{f_x}, \mathbf{f_u}, f_t) = \iiint \delta(\mathbf{x} - wt\mathbf{u}) R(|\mathbf{u}|/A) R(t/T) \cdot e^{-2\pi i(\mathbf{f_x}\cdot\mathbf{x} + \mathbf{f_u}\cdot\mathbf{u} + f_t t)} d\mathbf{x} d\mathbf{u} dt$$

$$= \iint R(|\mathbf{u}|/A) R(t/T)\, e^{-2\pi i((wt-s)\mathbf{f_x}\cdot\mathbf{u} + f_t t)} d\mathbf{u} dt, \quad (7)$$

where we have integrated the delta function over $\mathbf{x}$ and plugged $\mathbf{f_u} = -s\mathbf{f_x}$ from Eq. (5). Next, we integrate over $\mathbf{u}$. The 2D Fourier transform of a disc $R(|\mathbf{u}|/A)$ is a jinc function $(\pi A^2/4)\text{jinc}(\pi A|\mathbf{f_u}|)$ [4], where $\text{jinc}(z) = 2J_1(z)/z$, and $J_n(z)$ is the n-th order Bessel function of the first kind [16]. Since Eq. (7) has $(wt-s)\mathbf{f_x}$ as a frequency component for $\mathbf{u}$, we have

$$\hat{k} = \int \frac{\pi A^2}{4} \text{jinc}(\pi A(wt - s)|\mathbf{f_x}|) R(t/T) e^{-2\pi i f_t t} dt. \quad (8)$$

With infinite exposure assumption $T \to \infty$, Eq. (8) is the 1D Fourier transform of a jinc along the t axis. This produces the aforementioned fall-off along the f_t direction [3] (note $f_t = -\mathbf{m}\cdot\mathbf{f_x}$ as in Eq. (5)), which represents the deviation from perfect motion invariance.

Here we note that we can make $\hat{k}$ in Eq. (8) constant if we have a sinc function instead of the jinc, as the 1D Fourier transform of a sinc is a rect function. Now the question becomes as follows. *The jinc function resulted from the 2D Fourier transform of a disc $R(|\mathbf{u}|/A)$. What function produces a sinc function as the result of 2D Fourier transform?*

To answer this question, we take the inverse 2D Fourier transform of a sinc as

$$B_A(\mathbf{u}) = \int \frac{\pi A^2}{4} \text{sinc}(\pi A|\mathbf{f_u}|)\, e^{2\pi i \mathbf{u}\cdot\mathbf{f_u}} d\mathbf{u}. \quad (9)$$

Using polar coordinates as $\mathbf{f_u} = (f_r\cos\theta, f_r\sin\theta)$ with $f_r \equiv |\mathbf{f_u}|$,

$$B_A(\mathbf{u}) = \frac{\pi A^2}{4} \int_0^\infty \int_0^{2\pi} \text{sinc}(\pi A f_r) e^{2\pi i |\mathbf{u}| f_r \cos\theta} d\theta\, f_r\, df_r$$

$$= \frac{\pi A^2}{4} \int_0^\infty \text{sinc}(\pi A f_r) \cdot 2\pi J_0(2\pi|\mathbf{u}|f_r) f_r\, df_r$$

$$= \frac{\pi A}{2} \int_0^\infty \sin(\pi A f_r) J_0(2\pi|\mathbf{u}|f_r) df_r, \quad (10)$$

where we have used $J_0(z) = 2\pi \int_0^{2\pi} e^{iz\cos\theta} d\theta$ [16] and the definition of sinc. The solution of the integral in Eq. (10) can be found in [16], leading to (see the blue plot in Fig. 3)

$$B_A(\mathbf{u}) = \begin{cases} \dfrac{A}{2\sqrt{A^2 - 4|\mathbf{u}|^2}} & (|\mathbf{u}| < \frac{A}{2}) \\ \infty & (|\mathbf{u}| = \frac{A}{2}) \\ 0 & (|\mathbf{u}| > \frac{A}{2}) \end{cases}. \quad (11)$$

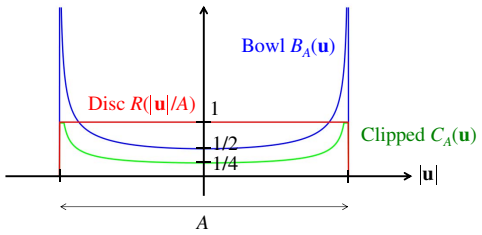


Figure 3. Profiles of circularly symmetric bokehs plotted with respect to the radius $|\mathbf{u}|$.

Hence, by using the following kernel k' instead of the conventional focus sweep kernel k in Eq. (2), one can achieve perfect 2D motion invariance (as well as depth invariance) in the limit of infinite exposure time.

$$k'(\mathbf{x}, \mathbf{u}, t) = \delta(\mathbf{x} - wt\mathbf{u}) B_A(\mathbf{u}) R(t/T). \quad (12)$$

Indeed, by taking the 5D Fourier transform of k', we can derive (see Appendix A for details)

$$|\hat{\phi}'_{s,\mathbf{m}}(\mathbf{f_x})|^2 = \begin{cases} \dfrac{\pi^2 A^2}{16 w^2 |\mathbf{f_x}|^2} \\ \quad (\text{for } |s| \le \frac{Tw}{2},\ |\mathbf{m}\cdot\mathbf{f_x}| \le \frac{Aw}{2}|\mathbf{f_x}|) \\ 0 \quad (\text{otherwise}) \end{cases} \quad (13)$$

which is constant irrespective of scene depth s and motion $\mathbf{m}$ for $|s| \le S/2$ $(= Tw/2)$ and $|\mathbf{m}| \le M/2$ $(= Aw/2)$ as can be seen in the blue plot in Fig. 2

In [3], two performance measures of joint defocus and motion deblurring are proposed. One is a high-frequency preservation measure defined as the worst-case squared MTF $\min_{s,\mathbf{m}} |\hat{\phi}_{s,\mathbf{m}}(\mathbf{f_x})|^2$ and the other is a PSF invariance measure defined as the ratio of the worst-case squared MTF to the best-case value $\frac{\min_{s,\mathbf{m}} |\hat{\phi}_{s,\mathbf{m}}(\mathbf{f_x})|^2}{\max_{s,\mathbf{m}} |\hat{\phi}_{s,\mathbf{m}}(\mathbf{f_x})|^2}$. Table 1 shows comparison between the conventional focus sweep kernel k in Eq. (2) and the improved kernel k' in Eq. (12) in terms of these performance measures. The improved focus sweep kernel has high-frequency preservation performance closer to the optimal than the conventional kernel, in addition to achieving motion invariance.

Table 1. Values of the high-frequency preservation measure and the PSF invariance measure for the conventional focus sweep kernel (with $S = Tw$ and $M = Aw/\sqrt{3}$ as in [3]) and the improved kernel (with $S = Tw$ and $M = Aw$ as derived above). Values in the parentheses show percentages to the upper bounds.

	High-freq. preserv.	PSF invariance
Upper bound	$\frac{2A^3T}{3SM\|\mathbf{f_x}\|^2}$	1
Conventional focus sweep kernel k	$\frac{2A^3T}{3\sqrt{3}SM\|\mathbf{f_x}\|^2}$ (57.7%)	$\frac{2}{3}$ (66.7%)
Improved focus sweep kernel k'	$\frac{\pi^2A^3T}{16SM\|\mathbf{f_x}\|^2}$ (92.5%)	1 (100%)

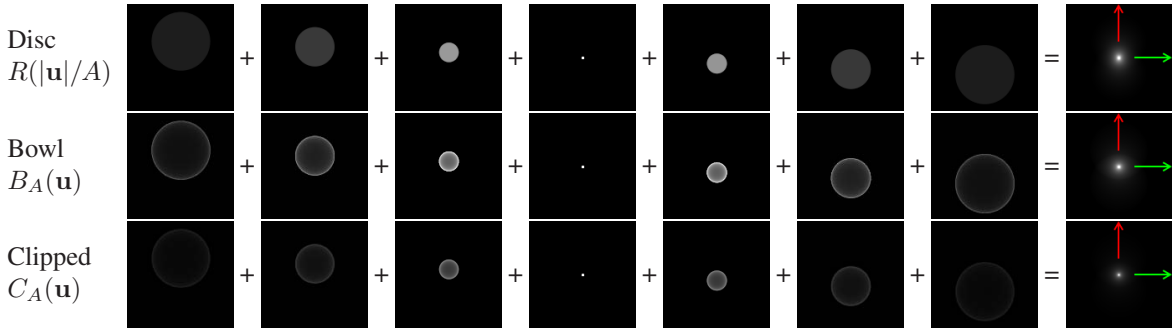


Figure 4. Log-intensity of bokehs (instantaneous PSFs) during focus sweep (seven images on the left for each row), and the resultant, time-integrated focus sweep PSFs (rightmost images). Top row: conventional focus sweep with a disc bokeh. Middle row: modified focus sweep with a bowl bokeh. Bottom row: modified focus sweep with a clipped bowl bokeh. Each row shows the case in which an object is moving vertically, and the bokeh center is shifting downward through time while changing its diameter. The vertical (red arrow) and horizontal (green arrow) profiles of the resultant focus sweep PSFs are shown in Fig. 5.

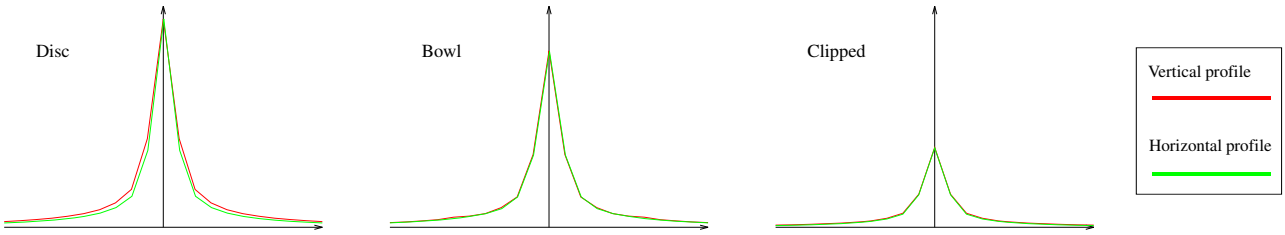


Figure 5. Focus sweep PSF profiles with a disc bokeh (left), a bowl bokeh (center), and a clipped bowl bokeh (right). Red and green plots show vertical and horizontal profiles, respectively, where the color corresponds to the same-colored arrows in Fig. 4. Note that the red plots are almost identical to (and thus hidden by) the green plots for the modified focus sweep (center and right).

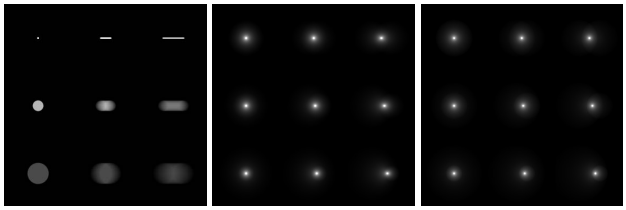
3.3. Modified Focus Sweep with a Custom Bokeh

The use of the new kernel of Eq. (12) means that one needs to distribute energy (or light rays) according to $B_A(\mathbf{u})$ inside the aperture rather than uniformly according to $R(|\mathbf{u}|/A)$. In other words, the defocus blur PSF of the lens used for focus sweep (which we call *bokeh* to distinguish it from the resultant focus sweep PSF) needs to have a bowl-like shape (the blue profile in Fig. 3) instead of a disc (the red profile). Although the bowl bokeh $B_A(\mathbf{u})$ has singularity at $|\mathbf{u}| = A/2$, it is integrable $\int B_A(\mathbf{u})d\mathbf{u} = \pi A^2/4 < \infty$, and hence it is physically realizable with finite resolution. An easy way to approximate the bowl bokeh is to place an attenuator at the lens aperture by sacrificing light, in which case large values have to be clipped as they cannot exceed $R(|\mathbf{u}|/A)$. We define such a clipped bokeh as

$$C_A(\mathbf{u}) = \min\{\alpha B_A(\mathbf{u}), R(|\mathbf{u}|/A)\}, \qquad (14)$$

where α is an attenuation coefficient. We find $\alpha = 1/2$ is a good compromise between the fidelity to the desirable profile and light efficiency (the green profile in Fig. 3).

While we conjecture that light efficient bowl bokehs may be realized with advanced and emerging optical elements such as phase plates and metamaterials [10], we leave the exploration of implementation as future work, and in what follows we evaluate the performance of the modified focus sweep with the ideal bowl bokeh $B_A(\mathbf{u})$ and its clipped version $C_A(\mathbf{u})$ by simulation.



Normal camera Focus sweep (disk) Focus sweep (bowl)

Figure 6. Simulated camera images of point light sources moving horizontally at three different speeds and at three different depths.

Using the ideal bowl bokeh, we can also confirm 2D motion invariance in the spatial domain. Let $\psi_b(\mathbf{x}) = \frac{4}{\pi b^2} B_b(\mathbf{x})$ denote the instantaneous defocus PSF with diameter b and unit volume. The focus sweep PSF is the integral of $\psi_b(\mathbf{x})$ with changing center position $\mathbf{x}$ and diameter b, and straightforward integration leads to (see Appendix B)

$$\phi'(\mathbf{x}) = \int_{-T/2}^{+T/2} \psi_{Aw|t|}(\mathbf{x} - \mathbf{m}t)dt \to \frac{1}{Aw|\mathbf{x}|} \qquad (15)$$

for $T \to \infty$, which does not depend on object motion $\mathbf{m}$.

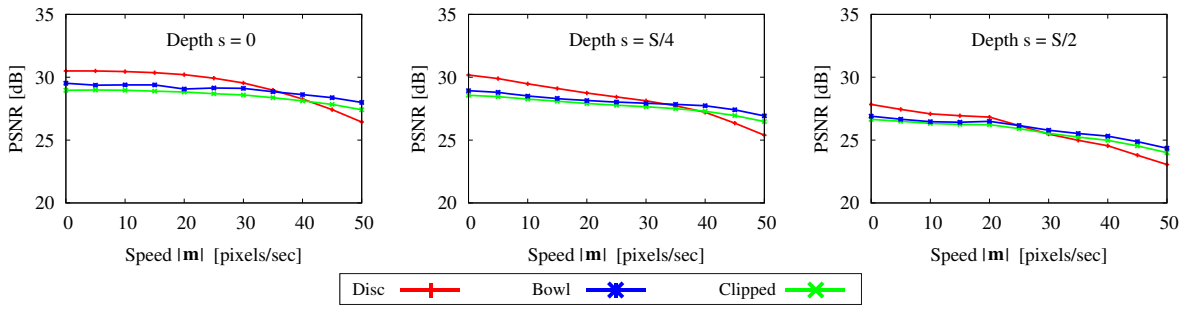


Figure 7. PSNR of deconvolution simulation results of focus sweep methods with different bokehs (disc, bowl, and clipped bowl). Scene depths $s = \{0, S/4, S/2\}$ and object speeds $|\mathbf{m}| \leq M/2$ are simulated, where $S = 1$ and $M = 100$.

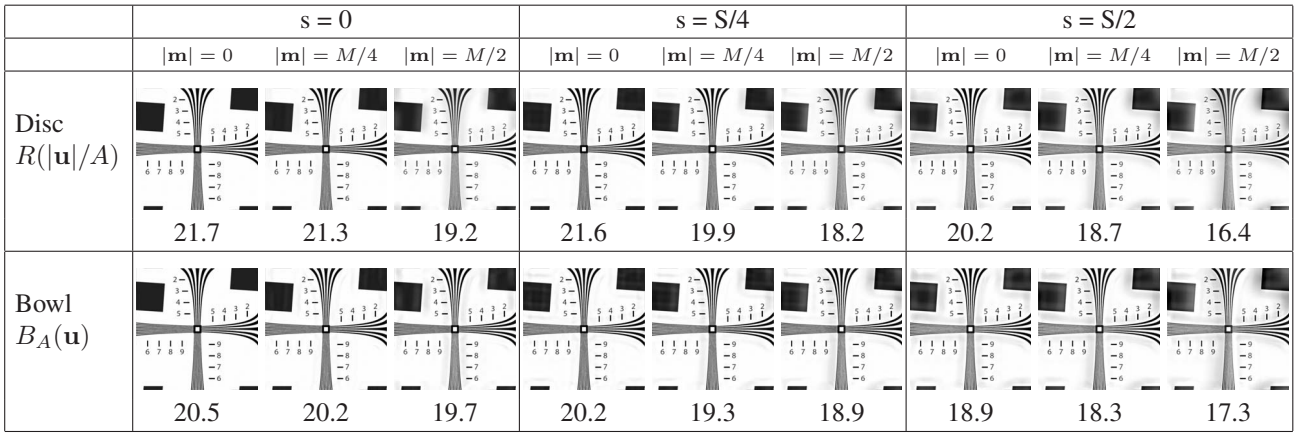


Figure 8. Deconvolved images and PSNR values from focus sweep simulation for various scene depths $s = \{0, S/4, S/2\}$ and speeds $|\mathbf{m}| = \{0, M/4, M/2\}$. Top row: conventional focus sweep with a disc bokeh. Bottom row: modified focus sweep with a bowl bokeh.

For a vertically moving scene point as shown in Fig. 4, a disc bokeh produces a vertically-elongated focus sweep PSF as can be seen in Fig. 5 (left) where its profile in the vertical direction (red) is wider than that in the horizontal direction (green). On the other hand, a bowl bokeh produces almost identical vertical and horizontal profiles, which remains also true for a clipped bowl bokeh, as shown in Fig. 5 (center and right). Fig. 6 shows more examples of focus sweep PSFs for varying object speeds and depths, along with normal camera PSFs for reference. The PSF elongation in the motion direction observed for the focus sweep with a disc bokeh is alleviated by the use of a bowl bokeh.

4. Evaluation

We conducted deblurring simulation for the conventional focus sweep with a disc bokeh and the modified focus sweep with bowl and clipped bokehs. We set $A = 100$ pixels, $T = 1$ sec, $S = 1$, and $M = 100$ pixels/sec, and simulated focus sweep PSFs with different bokehs for various object speeds and depths. We convolved a natural image with the PSF, added Gaussian noise with standard deviation 10^{-3} for [0, 1] pixel values, and computed the mean squared error (MSE) between the Wiener-deconvolved image and the original unblurred image. We repeated this process for several images and took the MSE average. In order to evaluate depth/motion invariance, we always used the "center" PSF corresponding to $s = 0$ and $\mathbf{m} = \mathbf{0}$ for deconvolution. As deconvolution with the center PSF can produce shifted images, we register the deconvolved image with the original image before computing the MSE.

Fig. 7 reports the simulation results in terms of PSNR = $-10 \log_{10}$(MSE). As can be seen, the performance of the conventional focus sweep gradually deteriorates for faster object motion. In contrast, the PSNR plot for the modified focus sweep with a bowl bokeh is flatter, producing better worst-case performance (i.e., minimum PSNR) at $|\mathbf{m}| = M/2$ $(= 50)$. The modified focus sweep with a clipped bowl bokeh performs slightly worse than with the ideal bowl bokeh due to light loss, but the degree of motion invariance almost remains the same as the ideal case. In practice, even the use of an ideal bowl bokeh cannot eliminate the deterioration for faster object motion and also for object depth away from the middle of the depth range $s = 0$. This is due to the use of a finite exposure time, known as a *tail-clipping effect* [11]. Please note that, as the modified

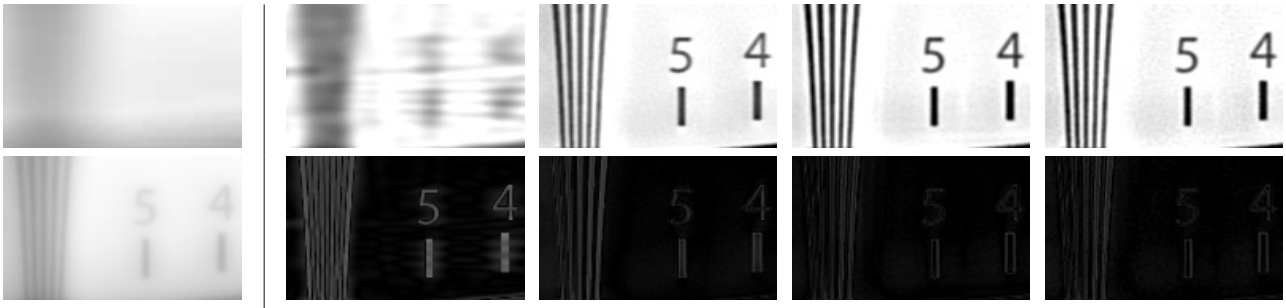


Figure 9. Magnified views of the deconvolution simulation results of a moving resolution chart in Fig. 8 with $s = S/2$ and $|\mathbf{m}| = M/2$. The leftmost column: simulated blurred images of the normal camera with a static lens (top) and of the focus sweep camera with a bowl bokeh (bottom, disk and clipped bowl bokehs produce similar images). The four columns on the right: deconvolution results (top) and their errors (bottom, differences from the ground truth unblurred image) of 1) the normal camera with a static lens, 2) the conventional focus sweep camera with a disc bokeh, 3) the modified focus sweep camera with a bowl bokeh, and of 4) the modified focus sweep camera with a clipped bowl bokeh.

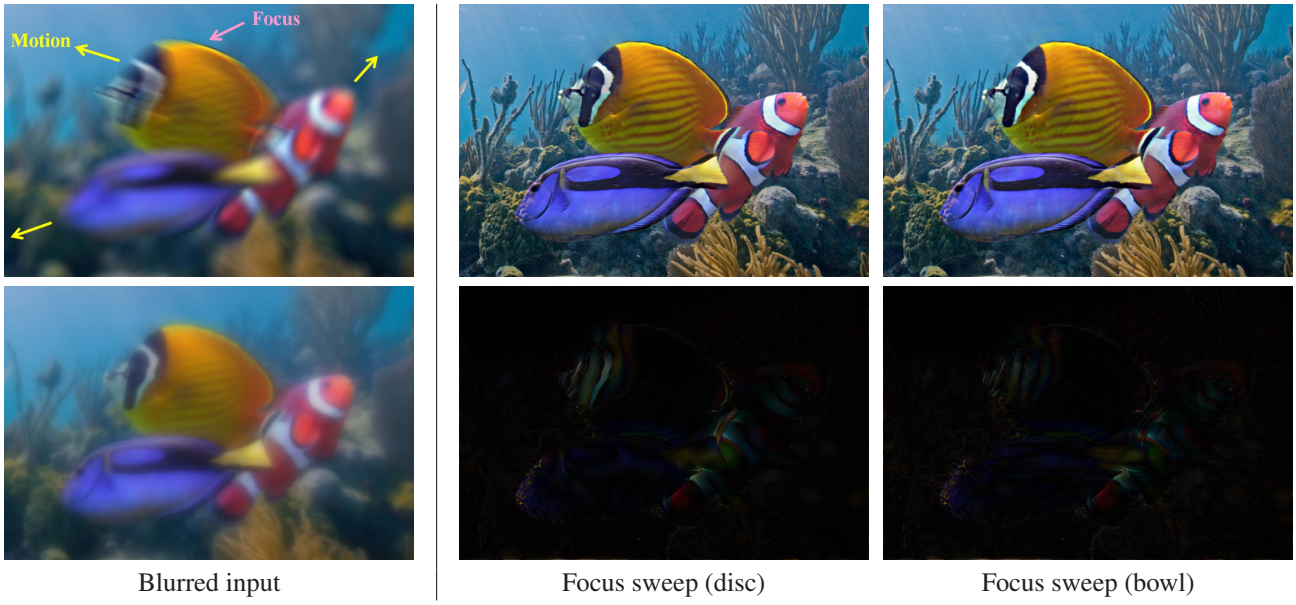


Figure 10. Deconvolution results of a simulated scene containing moving fish at different depths in front of a textured background of an ocean floor. The leftmost column: simulated blurred images of the normal camera with a static lens (top, focused on the yellow fish) and of the focus sweep camera with a bowl bokeh (bottom). The two columns on the right: deconvolution results (top) and their errors (bottom, differences from the ground truth unblurred image) of the focus sweep camera with a disc bokeh and a bowl bokeh.

focus sweep distributes the frequency power "budget" more evenly over the motion range, it comes with the cost of reduced PSNRs for slow object motion [11, 3]. Nevertheless, the modified focus sweep improves worst-case performance as dictated by the theory (see Table 1).

Fig. 8 shows simulated deconvolution results of a moving resolution chart at various depths. While the improvement of worst-case performance (at $s = S/2$ and $|\mathbf{m}| = M/2$) may not be visually significant, the modified focus sweep results in deblurred images with higher contrast as shown in the magnified views in Fig. 9. The use of a clipped bowl bokeh produces noisier images, but they still retain the overall contrast, providing better reconstructions than the conventional focus sweep.

Fig. 10 shows a simulated scene of moving fish. It consists of four depth layers (an ocean floor background and three fish), and is rendered with ray-tracing to simulate defocus and motion blur. Hence, the rendered images contain blur that cannot be modeled as simple convolution at occlusions. Nevertheless, as the focus sweep PSF remains nearly uniform over the image, visually pleasing images are recovered using Wiener deconvolution, with the focus sweep with a bowl bokeh producing better contrast than with a disc bokeh, as can be seen around the face of the yellow fish.

5. Conclusion

Through a time-varying light field analysis of the focus sweep PSF, this paper has shown that perfect 2D motion invariance is possible in theory, in the limit of infinite exposure time, by using a bowl-shaped lens bokeh instead of a standard disc bokeh for focus sweep. We have also verified that the use of a bowl bokeh improves motion invariance also in practice, and showed that it produces better worst-case performance than the conventional focus sweep. Although the improvement is small and may not justify the cost of designing a custom bokeh at present, we hope that emerging optics technologies will minimize such concerns in the future.

Our primary goal in this paper is to provide an analysis to answer the question of whether or not the gap between the theoretical optimum and the near-optimum achieved by the conventional focus sweep can be further reduced. While uniqueness of the 2D motion-invariant kernel and existence of kernels that also achieve optimal high-frequency preservation are yet to be investigated, we believe that the analysis presented in the paper provides further theoretical support not only for motion invariance of focus sweep but also for joint defocus and motion deblurring in general, upon which follow-on work can build.

Acknowledgments

The author would like to thank Matthew Hirsch, Gordon Wetzstein, and the anonymous reviewers for their valuable comments and suggestions.

References

[1] A. Agrawal and R. Raskar. Optimal single image capture for motion deblurring. In *CVPR*, pages 2560–2567, 2009. 1

[2] J. Baek. Transfer efficiency and depth invariance in computational cameras. In *ICCP*, pages 1–8, 2010. 1

[3] Y. Bando, H. Holtzman, and R. Raskar. Near-invariant blur for depth and 2D motion via time-varying light field analysis. *ACM Trans. Gr.*, 32(2):13:1–13:15, 2013. 1, 2, 3, 6

[4] M. Born and E. Wolf. *Principles of Optics, sixth (corrected) edition*. Pergamon Press, 1984. 3

[5] J.-X. Chai, X. Tong, S.-C. Chan, and H.-Y. Shum. Plenoptic sampling. In *Proc. SIGGRAPH 2000*, pages 307–318, 2000. 2

[6] O. Cossairt and S. Nayar. Spectral focal sweep: Extended depth of field from chromatic aberrations. In *ICCP*, pages 1–8, 2010. 1

[7] O. Cossairt, C. Zhou, and S. K. Nayar. Diffusion coded photography for extended depth of field. *ACM Trans. Gr.*, 29(4):31:1–31:10, 2010. 1

[8] E. R. Dowski and W. T. Cathey. Extended depth of field through wave-front coding. *Applied Optics*, 34(11):1859–1866, 1995. 1

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

[10] J. Hunt, T. Driscoll, A. Mrozack, G. Lipworth, M. Reynolds, D. Brady, and D. R. Smith. Metamaterial apertures for computational imaging. *Science*, 339(6117):310–313, 2013. 4

[11] A. Levin, P. Sand, T. S. Cho, F. Durand, and W. T. Freeman. Motion-invariant photography. *ACM Trans. Gr.*, 27(3):71:1–71:9, 2008. 1, 2, 5, 6

[12] M. Levoy and P. Hanrahan. Light field rendering. In *Proc. ACM SIGGRAPH 96*, pages 31–42, 1996. 2

[13] H. Nagahara, S. Kuthirummal, C. Zhou, and S. K. Nayar. Flexible depth of field photography. In *ECCV*, pages 60–73, 2008. 1

[14] R. Ng. Fourier slice photography. *ACM Trans. Gr.*, 24(3):735–744, 2005. 2

[15] R. Raskar, A. Agrawal, and J. Tumblin. Coded exposure photography: motion deblurring using fluttered shutter. *ACM Trans. Gr.*, 25(3):795–804, 2006. 1

[16] G. N. Watson. *A treatise on the theory of Bessel functions*. Cambridge University Press, 1922. 3

[17] D. Znamenskiy, H. Schmeitz, and R. Muijs. Motion invariant imaging by means of focal sweep. In *IEEE International Conference on Consumer Electronics*, pages 91–92, 2011. 1, 2

Appendix

A. MTF of the Improved Focus Sweep Kernel

Here we show a derivation of Eq. (13), the MTF of the improved focus sweep kernel.

We take the 5D Fourier transform of the improved focus sweep kernel in Eq. (12). First, we integrate the delta function over $\mathbf{x}$ and obtain

$$\begin{aligned}\hat{k}'(\mathbf{f_x},\mathbf{f_u},f_t) &= \iiint \delta(\mathbf{x}-wt\mathbf{u})B_A(\mathbf{u})R(t/T) \\ &\qquad \cdot e^{-2\pi i(\mathbf{f_x}\cdot\mathbf{x}+\mathbf{f_u}\cdot\mathbf{u}+f_t t)}d\mathbf{x}d\mathbf{u}dt \\ &= \iint B_A(\mathbf{u})R(t/T)e^{-2\pi i(\mathbf{f_x}\cdot(wt\mathbf{u})+\mathbf{f_u}\cdot\mathbf{u}+f_t t)}d\mathbf{u}dt \\ &= \iint B_A(\mathbf{u})R(t/T)e^{-2\pi i((wt-s)\mathbf{f_x}\cdot\mathbf{u}+f_t t)}d\mathbf{u}dt, \quad \text{(A.1)}\end{aligned}$$

where for the last line we have substituted Eq. (5) for $\mathbf{f_u}$. Next, we integrate over $\mathbf{u}$. Since we have shown that the 2D Fourier transform of a bowl $B_A(\mathbf{u})$ is a sinc: $(\pi A^2/4)\text{sinc}(\pi A|\mathbf{f_u}|)$, and Eq. (A.1) has $(wt-s)\mathbf{f_x}$ as a frequency component for $\mathbf{u}$, we have

$$\hat{k}' = \int \frac{\pi A^2}{4}\text{sinc}(\pi A(wt-s)|\mathbf{f_x}|)R(t/T)e^{-2\pi i f_t t}dt. \quad \text{(A.2)}$$

Finally, we integrate over t. For the moment, we omit $R(t/T)$ by assuming infinite exposure time. We rearrange

Eq. (A.2) with change of variable as $t' = t - s/w$ and obtain

$$\hat{k}' = \frac{\pi A^2}{4} e^{-2\pi i f_t s/w} \int \text{sinc}(\pi A w |\mathbf{f_x}| t') e^{-2\pi i f_t t'} dt'. \tag{A.3}$$

This amounts to the 1D Fourier transform of a sinc. As the Fourier transform of $\text{sinc}(at)$ with respect to t is given as $(\pi/a)R(\pi f_t/a)$, applying this to Eq. (A.3) and taking the magnitude leads to

$$|\hat{k}'|^2 = \begin{cases} \dfrac{\pi^2 A^2}{16 w^2 |\mathbf{f_x}|^2} & \left(|f_t| \le \dfrac{A}{2} w |\mathbf{f_x}|\right) \\ 0 & \text{(otherwise)} \end{cases} \tag{A.4}$$

With finite exposure time, Eq. (A.3) gets convolved by the Fourier transform of $R(t/T)$, which is $T\text{sinc}(\pi T f_t)$, in the f_t axis. Since convolution by $\text{sinc}(\pi T f_t)$ cancels out sinusoids with higher frequencies than $T/2$, and since Eq. (A.3) has the sinusoid term $e^{-2\pi i f_t s/w}$, the additional condition for Eq. (A.4) to be non-zero is given as $|s| \le (T/2)w$. Plugging Eq. (5) for f_t into Eq. (A.4) leads to Eq. (13).

B. Proof of Perfect 2D Motion Invariance

Here we prove perfect 2D motion invariance of the modified focus sweep by showing a derivation of its PSF given in Eq. (15).

We start from the left hand side of Eq. (15).

$$\phi'(\mathbf{x}) = \int_{-T/2}^{+T/2} \psi_{Aw|t|}(\mathbf{x} - \mathbf{m}t) dt. \tag{B.1}$$

For $|\mathbf{m}| < M/2$ $(= Aw/2)$, the above equation can be written as

$$\phi'(\mathbf{x}) = \int_{-T/2}^{t_0} \frac{1}{\pi A w |t| \sqrt{q(t)}} dt + \int_{t_1}^{+T/2} \frac{1}{\pi A w |t| \sqrt{q(t)}} dt, \tag{B.2}$$

where $q(t) = (Awt/2)^2 - |\mathbf{x} - \mathbf{m}t|^2$, and t_0 and t_1 are the roots of $q(t) = 0$. We order them such that $t_0 \le t_1$, and we assume T is large enough to satisfy $-T/2 < t_0$ and $t_1 < +T/2$. If we further set as

$$\begin{aligned} q(t) &= (A^2 w^2/4 - |\mathbf{m}|^2) t^2 + 2(\mathbf{m} \cdot \mathbf{x}) t - |\mathbf{x}|^2 \\ &\equiv at^2 + bt + c, \end{aligned} \tag{B.3}$$

where $a \equiv A^2 w^2/4 - |\mathbf{m}|^2 > 0$, $b \equiv 2(\mathbf{m} \cdot \mathbf{x})$, and $c \equiv -|\mathbf{x}|^2 \le 0$, we can write the roots as

$$\begin{aligned} \{t_0, t_1\} &= \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \\ &= \frac{-(\mathbf{m} \cdot \mathbf{x}) \pm \sqrt{(\mathbf{m} \cdot \mathbf{x})^2 + (A^2 w^2/4 - |\mathbf{m}|^2)|\mathbf{x}|^2}}{(A^2 w^2/4 - |\mathbf{m}|^2)}. \end{aligned} \tag{B.4}$$

Since $a > 0$, $c \le 0$, and therefore $b^2 - 4ac \ge b^2 \ge 0$, existence of the real roots t_0 and t_1 is guaranteed, and we also have $t_0 \le 0$ and $t_1 \ge 0$. Now we can remove the abs operations in Eq. (B.2) as

$$\phi'(\mathbf{x}) = \int_{-T/2}^{t_0} -\frac{1}{\pi A w t \sqrt{q(t)}} dt + \int_{t_1}^{+T/2} \frac{1}{\pi A w t \sqrt{q(t)}} dt. \tag{B.5}$$

Here we apply the following equation

$$\int \frac{1}{t\sqrt{at^2 + bt + c}} dt = \frac{1}{\sqrt{-c}} \sin^{-1}\left(\frac{bt + 2c}{|t|\sqrt{b^2 - 4ac}}\right). \tag{B.6}$$

Then,

$$\begin{aligned} \phi'(\mathbf{x}) = &\left[-\frac{1}{\pi A w |\mathbf{x}|} \sin^{-1}\left(\frac{bt + 2c}{|t|\sqrt{b^2 - 4ac}}\right)\right]_{-T/2}^{t_0} \\ &+ \left[\frac{1}{\pi A w |\mathbf{x}|} \sin^{-1}\left(\frac{bt + 2c}{|t|\sqrt{b^2 - 4ac}}\right)\right]_{t_1}^{+T/2}. \end{aligned} \tag{B.7}$$

By simple substitution, we can see that

$$\frac{bt_0 + 2c}{|t_0|\sqrt{b^2 - 4ac}} = \frac{\frac{-b^2 - b\sqrt{b^2 - 4ac} + 4ac}{2a}}{-\frac{-b\sqrt{b^2 - 4ac} - (b^2 - 4ac)}{2a}} = -1, \tag{B.8}$$

$$\frac{bt_1 + 2c}{|t_1|\sqrt{b^2 - 4ac}} = \frac{\frac{-b^2 + b\sqrt{b^2 - 4ac} + 4ac}{2a}}{\frac{-b\sqrt{b^2 - 4ac} + (b^2 - 4ac)}{2a}} = -1, \tag{B.9}$$

and thus

$$\begin{aligned} \phi'(\mathbf{x}) &= \frac{1}{\pi A w |\mathbf{x}|} \left[-\sin^{-1}(-1) + \sin^{-1}\left(\frac{-bT + 4c}{T\sqrt{b^2 - 4ac}}\right)\right. \\ &\qquad \left. + \sin^{-1}\left(\frac{bT + 4c}{T\sqrt{b^2 - 4ac}}\right) - \sin^{-1}(-1)\right] \\ &= \frac{1}{\pi A w |\mathbf{x}|} \left[\frac{\pi}{2} + \sin^{-1}\left(\frac{-b + 4c/T}{\sqrt{b^2 - 4ac}}\right)\right. \\ &\qquad \left. + \sin^{-1}\left(\frac{b + 4c/T}{\sqrt{b^2 - 4ac}}\right) + \frac{\pi}{2}\right] \\ &= \frac{1}{Aw|\mathbf{x}|} + \frac{1}{\pi A w |\mathbf{x}|} \left[\sin^{-1}\left(\frac{b + 4c/T}{\sqrt{b^2 - 4ac}}\right)\right. \\ &\qquad \left. - \sin^{-1}\left(\frac{b - 4c/T}{\sqrt{b^2 - 4ac}}\right)\right], \end{aligned} \tag{B.10}$$

as $\sin^{-1}(\cdot)$ is an odd function. Now we can see that

$$\phi'(\mathbf{x}) \to \frac{1}{Aw|\mathbf{x}|} \tag{B.11}$$

for $T \to \infty$.