

Motion Deblurring from a Single Image using Circular Sensor Motion

Yosuke Bando^{†,§}, Bing-Yu Chen[‡], and Tomoyuki Nishita[§]

[†]TOSHIBA Corporation

[§]The University of Tokyo

[‡]National Taiwan University

Abstract

Image blur caused by object motion attenuates high frequency content of images, making post-capture deblurring an ill-posed problem. The recoverable frequency band quickly becomes narrower for faster object motion as high frequencies are severely attenuated and virtually lost. This paper proposes to translate a camera sensor circularly about the optical axis during exposure, so that high frequencies can be preserved for a wide range of in-plane linear object motion in any direction within some predetermined speed. That is, although no object may be photographed sharply at capture time, differently moving objects captured in a single image can be deconvolved with similar quality. In addition, circular sensor motion is shown to facilitate blur estimation thanks to distinct frequency zero patterns of the resulting motion blur point-spread functions. An analysis of the frequency characteristics of circular sensor motion in relation to linear object motion is presented, along with deconvolution results for photographs captured with a prototype camera.

Categories and Subject Descriptors (according to ACM CCS): I.4.1 [Image Processing and Computer Vision]: Digitization and Image Capture— I.4.3 [Image Processing and Computer Vision]: Enhancement—Sharpening and deblurring

1. Introduction

Motion blur often spoils photographs by losing image sharpness. As motion blur attenuates high frequency content of images, motion deblurring is an ill-posed problem and often comes with noise amplification and ringing artifacts [BLM90]. Although image deconvolution techniques advance continuously to tackle this problem, motion deblurring is still challenging since the recoverable frequency band easily becomes narrow for fast object motion as high frequencies are severely attenuated and virtually lost.

A simple countermeasure, called *follow shot*, can capture sharp images of a moving object as if it were static by panning a camera to track the object during exposure. However, there are cases where follow shot is not effective: 1) when object motion is unpredictable; 2) when there are multiple

objects with different motion. This is because follow shot favors particular motion one has chosen to track, as much as a static camera favors “motion” at the speed of zero (i.e., static objects): objects moving differently from favored motion degrade.

This paper explores a camera hardware-assisted approach to single-shot motion deblurring that simultaneously preserves high frequency image content for different object motion. Under the assumption that object motion is in-plane linear (having arbitrary 2D directions) within some predetermined speed, we propose to translate a camera sensor circularly about the optical axis during exposure, so that the camera partially “follow-shots” various object motion. As a result, differently moving objects can be deconvolved with similar quality.

Our work is inspired by Levin *et al.* [LSC*08], who proved that constantly accelerating 1D sensor motion can render motion blur *invariant* to 1D linear object motion (e.g., horizontal motion), and showed that this sensor motion evenly distributes the fixed frequency “budget” to different

[†] yosuke1.bando@toshiba.co.jp

[‡] robin@ntu.edu.tw

[§] {ybando, nis}@nis-lab.is.s.u-tokyo.ac.jp

object speeds. Similarly to Cho *et al.* [CLDF10], we intend to extend Levin *et al.*'s budgeting argument to 2D (i.e., in-plane) linear object motion by sacrificing motion-invariance. However, unlike Cho *et al.*'s and other researchers' multi-shot approaches [CLDF10, RAP05, YSQS07, AXR09], in this paper we would like to seek a single-shot solution because increasing the number of exposures may incur other issues such as capture time overhead/delay between multiple exposures, signal-to-noise ratio degradation if shorter exposure time per shot is used, or violation of the linear object motion model if total exposure time becomes longer.

By losing motion-invariance, we inevitably reintroduce an issue inherent to the classical motion deblurring problem, which [LSC*08] resolved for 1D motion; we need to locally estimate a *point-spread function* (PSF) of motion blur as it depends on object motion. Fortunately, although we rely on user intervention to segment images into different moving objects, we can show that PSF discriminability for each moving object is higher for the circular sensor motion camera than the previous single-shot image capture strategies, thanks to distinct frequency zero patterns of the PSFs.

Although we cannot guarantee worst-case deblurring performance as the PSFs have frequency zeros, circular sensor motion can be shown to provide $2/\pi$ ($\approx 64\%$) of the optimal frequency bound on an average. We show deconvolution results for simulated images as well as real photographs captured by our prototype camera, and also demonstrate other advantages of circular sensor motion: 1) motion blurred objects in an image are recognizable (e.g., text is readable) even without deconvolution; 2) the circular motion strategy has no 180° motion ambiguity in PSF estimation; it can distinguish rightward object motion from leftward one.

2. Related Work

Capture time approach: Motion blur can be reduced using short exposure time, but signal-to-noise ratio worsens due to loss of light. Recent cameras are equipped with image stabilization hardware that shifts the lens or sensor to compensate for camera motion acquired from a gyroscope. This is effective for preventing camera shake blur but not for object motion blur. Ben-Ezra and Nayar [BEN04] acquired camera motion from a low resolution video camera attached to a main camera, which was used for camera shake removal for the main camera. Tai *et al.* [TDBL08] extended their approach to handle videos with non-uniform blur. Joshi *et al.* [JKZS10] used a 3-axis accelerometer and gyroscopes to guide camera shake removal.

Raskar *et al.* [RAT06] developed a *coded exposure* technique to prevent attenuation of high frequencies due to motion blur at capture time by opening and closing the shutter during exposure according to a pseudo-random binary code. The method was extended to be capable of PSF estimation [AX09] and to handle non-uniform/nonlinear blur [TKLS10, DMY10].

The trade-offs among the coded exposure photography, motion-invariant photography [LSC*08], and ours are summarized as follows (refer also to [AR09] for detailed comparison between the coded exposure and motion-invariant strategies). A static camera can capture static objects perfectly, but high frequencies will be rapidly lost as object motion gets faster. The coded exposure strategy significantly reduces this loss of frequencies. The motion-invariant strategy best preserves high frequencies for 1D (horizontal) object motion up to the predetermined speed, denoted by S , but it does not generalize to other motion directions. The circular motion strategy can treat any direction, and it achieves better high frequency preservation for target object speed S than the coded exposure strategy. Similar to the motion-invariant strategy, the circular motion strategy degrades static scene parts due to sensor motion, but it can partially track moving objects so that they are recognizable even before deconvolution. Unlike the other strategies, the circular motion strategy has no 180° motion ambiguity in PSF estimation. These trade-offs will be explained in more detail and demonstrated in the following sections.

Cho *et al.* [CLDF10] proposed a two-shot approach with the motion-invariant strategy aimed to two orthogonal directions (i.e., horizontal and vertical). In contrast, this paper pursues a single shot approach. Direct comparison with multi-shot approaches requires elaborate modeling of capture time overhead and noise and is out of the scope of this paper, but we will present some observation in Sec. 4.

Post-capture approach (PSF estimation and image deconvolution): This field has a large body of literature, and we refer the readers to [KH96] for the early work. Recently, significant advancement was brought forth by the incorporation of sophisticated regularization schemes and by extending the range of target blur to non-uniform and/or large ones [FSH*06, Lev06, Jia07, SXJ07, SJA08, YSQS08, JSK08, CL09, XJ10, KTF11]. Some researchers used multiple images [RAP05, YSQS07, AXR09, ZGS10], some of which use different exposure times or flash/no-flash image pairs.

Other applications of sensor motion: Some researchers proposed to move sensors for different purposes. Ben-Ezra *et al.* [BEZN05] moved the sensor by a fraction of a pixel size between exposures for video super-resolution. Mohan *et al.* [MLHR09] moved the lens and sensor to deliberately introduce motion blur that acts like defocus blur. Nagahara *et al.* [NKZN08] moved the sensor along the optical axis to make defocus blur depth-invariant.

3. Circular Image Integration

Fig. 1(a) shows the proposed motion of a camera image sensor. We *translate* the sensor along a circle perpendicular to the optical axis while *keeping its orientation*. We use the phrase "circular motion" to emphasize that we do not *rotate* the sensor itself.

During exposure time $t \in [-T, +T]$, the sensor undergoes one revolution with constant angular velocity $\omega = \pi/T$. Letting the radius of circular motion be R , the sensor moves along the circle with constant speed $R\omega$, which corresponds to the target object speed S in the image space. The corresponding object speed in the world space (i.e., actual speed in a scene) is determined by the camera optics and the distance to the object from the camera. Given exposure time $2T$ and the target object speed S , the appropriate radius is therefore $R = ST/\pi$. Taking an xy plane on the sensor, the sensor motion goes through a spiral in the xyt space-time volume as shown in red in Fig. 1(b).

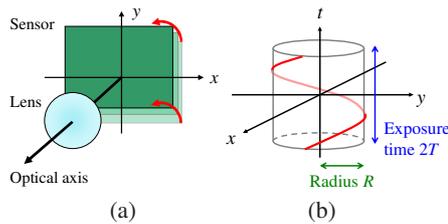


Figure 1: Circular sensor motion. (a) The sensor is translated circularly about the optical axis. (b) Sensor motion trajectory (red curve) in the space-time volume.

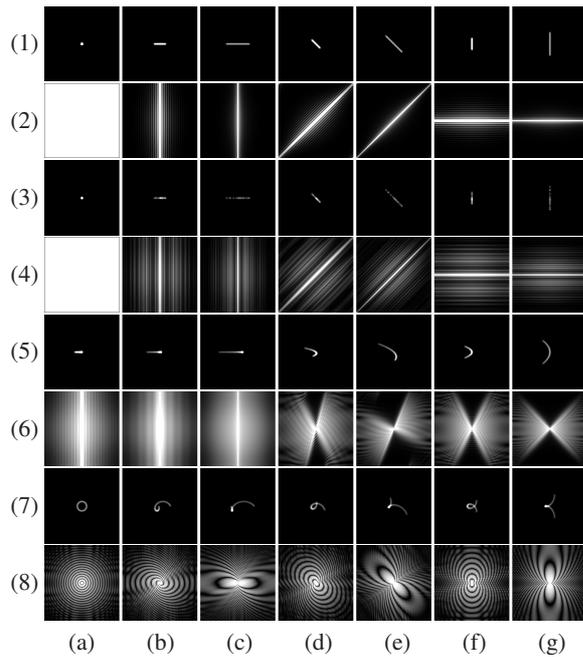


Figure 2: Motion blur PSFs and their corresponding log power spectra. Rows: (1) PSFs and (2) power spectra for a static camera. (3)(4) Coded exposure camera. (5)(6) Motion-invariant camera. (7)(8) Circular motion camera. Columns: (a) Static object. (b)(c) Horizontal object motion at different speeds. (d)(e) Oblique object motion. (f)(g) Vertical motion.

Fig. 2 shows simulated motion blur PSFs and their power spectra of various object motions observed from a static camera, the coded exposure camera [RAT06], the motion-invariant camera [LSC*08], and our circular motion camera. While the power spectrum of a static object for a static camera is perfectly broadband, those of moving objects become quickly narrowband as the object speed increases. The coded exposure camera makes power spectra broadband at the cost of losing light blocked by the shutter, but the tendency of bandwidth narrowing for faster motion remains. The motion-invariant camera produces similarly broadband power spectra for horizontal motions (they are not completely identical due to the *tail clipping* effect [LSC*08]), but vertical frequencies are sacrificed as motion direction deviates from horizontal. The circular motion camera produces power spectra that extend to high frequency regions in all cases. Although they have striped frequency zeros, these zeros facilitate PSF estimation as described in Sec. 5.

4. Analysis

Levin *et al.* [LSC*08] proved that constantly accelerating 1D sensor motion is the only sensor motion that makes PSF invariant to 1D linear object motion. From this finding we can see that there is no sensor motion that makes PSF invariant to 2D linear object motion. Hence, we must abandon motion-invariance, and we seek to extend Levin *et al.*'s another finding that their sensor motion evenly and nearly optimally distributes the fixed frequency “budget” to different object speeds.

The intuitive explanation for optimality of constant camera acceleration for 1D case is as follows. Fig. 3(a) shows the range of speed $[-S, +S]$ that must be taken care of. We can cover the entire range by accelerating a camera beginning at speed $-S$ until it reaches $+S$. The camera tracks every speed at one moment during exposure. By extending to 2D, the range of *velocity* (speed + direction) we must cover becomes a disc as shown in green in Fig. 3(b). We are no longer able to fill the entire disc by a finite sensor motion path, and we opt to trace only the circumference of the disc (shown in blue), which can be achieved by moving a sensor circularly. The reasons for doing so are threefold.

1. It makes theoretical analysis easier. Although full frequency analysis of 3D xyt space-time is difficult, we were able to draw some insights into frequency characteristics of circular sensor motion.
2. Tracing the circumference alone can be shown to deal with velocity in the interior of the disc fairly well.
3. It makes implementation of camera hardware easier.

Multi-revolution and multi-shot approaches: As for Reason 2 above, to further treat different object speeds evenly, one can consider sampling the interior of the velocity disc by a set of concentric circles. However, this does not bring in significant improvement of PSF power spectra, since the

phases of the Fourier transform of multiple circular motions cancel each other when superimposed, resulting in a set of power spectra as shown in Fig. 4(1), which is qualitatively similar to the one shown in the bottom row of Fig. 2. That is, for multiple PSFs whose Fourier transforms are given as $F_i(f_x, f_y)$ ($i = 0, 1, \dots$), the combined power spectrum $|\sum_i F_i(f_x, f_y)|^2$ can be zero for some frequency (a, b) even if $|F_i(a, b)|^2 > 0$ for some or all i . If multiple shots were allowed, the phase cancellation would not occur, and the combined power spectrum would be $\sum_i |F_i(f_x, f_y)|^2$, meaning that frequency zeros of one PSF could be filled with non-zero frequencies of the other PSFs [AXR09]. Therefore, whether or not to take multi-shot approaches depends on a cost model for multiple shots (if there is no cost, taking many images is preferable, for example), which is left as future work. For a single shot approach, moving a sensor in two orthogonal directions as in [CLDF10] during exposure produces power spectra shown in Fig. 4(2), which have tendency of bandwidth narrowing for faster object motion.

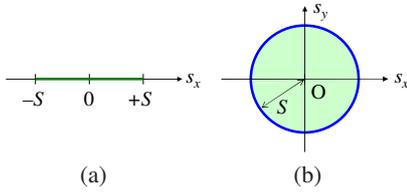


Figure 3: The range of velocity (s_x, s_y) that needs to be covered for (a) 1D case and (b) 2D case (shown in green). We trace only the circumference of the disc (shown in blue).

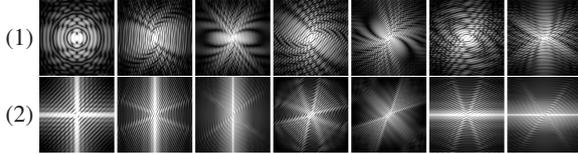


Figure 4: Power spectra of the motion blur PSFs from (1) two-revolution circular sensor motion and (2) horizontal sensor motion followed by vertical one. The order of columns is the same as in Fig. 2.

4.1. Frequency Budgeting

Now we review the frequency budgeting argument of [LSC*08] for the case of 2D object motion.

We consider a camera path in the xyt space-time volume.

$$p(\mathbf{x}, t) = \begin{cases} \delta(\mathbf{x} - \mathbf{m}(t)) & \text{for } t \in [-T, +T] \\ 0 & \text{otherwise} \end{cases}, \quad (1)$$

where $\mathbf{x} = (x, y)$, $\mathbf{m}(t)$ specifies the camera position at time t , and $\delta(\cdot)$ is a delta function. We would like to consider its 3D Fourier transform, denoted by \hat{p} :

$$\hat{p}(\mathbf{f}, f_t) = \int_{\Omega} \int_{-T}^{+T} \delta(\mathbf{x} - \mathbf{m}(t)) e^{-2\pi i(\mathbf{f} \cdot \mathbf{x} + f_t t)} dt d\mathbf{x}, \quad (2)$$

where $\mathbf{f} = (f_x, f_y)$ is a 2D spatial frequency, f_t is a temporal frequency, and Ω spans the entire xy plane.

It can be shown that the 2D Fourier transform of a motion blur PSF for object velocity \mathbf{v} is a 2D slice of $\hat{p}(\mathbf{f}, f_t)$ along the plane of $f_t = -\mathbf{v} \cdot \mathbf{f} = -s_x f_x - s_y f_y$ (Fourier projection-slice theorem [Bra65]). Therefore, given a maximum speed S , the volume in the 3D $f_x f_y f_t$ frequency domain that these slices can pass through is confined to the outside of the cone as $|f_t| \leq S|\mathbf{f}|$, called the *wedge of revolution* in [CLDF10], as shown in blue in Fig. 5(a). We would like $|\hat{p}(\mathbf{f}, f_t)|$ to have as large value as possible within this volume, so that motion blur PSFs up to S have large power spectra. However, the budget is exactly $2T$ along each vertical line $\mathbf{f} = \mathbf{c}$ (the line shown in red and green in Fig. 5(a)) for any given spatial frequency \mathbf{c} : i.e., $\int |\hat{p}(\mathbf{c}, f_t)|^2 df_t = 2T$ [LSC*08].

To assign the $2T$ budget so that any 2D linear object motion below S has a similar amount of PSF spectral power, we consider the following two criteria.

Effectiveness: The budget should be assigned as much as possible within the line segment of $f_t \in [-S|\mathbf{c}|, +S|\mathbf{c}|]$ which is shown in red in Fig. 5(a). In other words, we would like to avoid assigning the budget to the other portions of the line (shown in green in Fig. 5(a)) as they correspond to object speeds beyond S and the budget will be wasted. Because the budget is exactly $2T$ unless we close the shutter during exposure, less assignment to some portion means more assignment to the other.

Uniformity: The budget should be distributed evenly across the line segment, so that every object motion PSF has an equal amount of spectral power.

Therefore, optimal assignment in which both effectiveness and uniformity are perfect gives $T/S|\mathbf{c}|$ to each point on the line segment.

4.2. Spectrum of Circular Sensor Motion

Now we take the 3D Fourier transform of the circular sensor motion $\mathbf{m}(t) = (R \cos \omega t, R \sin \omega t)$, a spiral in the xyt space-time as shown in Fig. 1(b). By integrating Eq. (2) with respect to t , we obtain:

$$\hat{p}(\mathbf{f}, f_t) = \int_{\Omega} \left(\frac{\delta(|\mathbf{x}| - R)}{R\omega} e^{-2\pi i f_t \mathbf{m}^{-1}(\mathbf{x})} \right) e^{-2\pi i \mathbf{f} \cdot \mathbf{x}} d\mathbf{x}, \quad (3)$$

since the integrand is non-zero only at $|\mathbf{x}| = R$ and at $t = \mathbf{m}^{-1}(\mathbf{x})$. Jacobian $|d\mathbf{m}(t)/dt| = R\omega$ is introduced in the denominator. By using polar coordinates as $x = r \cos \theta$ and $y = r \sin \theta$,

$$\hat{p}(\mathbf{f}, f_t) = \int_{\Omega} \left(\frac{\delta(r - R)}{R\omega} e^{-2\pi i f_t \theta / \omega} \right) e^{-2\pi i \mathbf{f} \cdot \mathbf{x}} d\mathbf{x}. \quad (4)$$

This is a hard-to-integrate expression, but we can proceed if we focus on a set of discrete f_t slices where $k = 2\pi f_t / \omega$ is an integer as shown in Fig. 5(b), as (see Appendix A):

$$|\hat{p}(\mathbf{f}, f_t)|^2 = 4T^2 J_k^2(2\pi R|\mathbf{f}|), \quad (5)$$

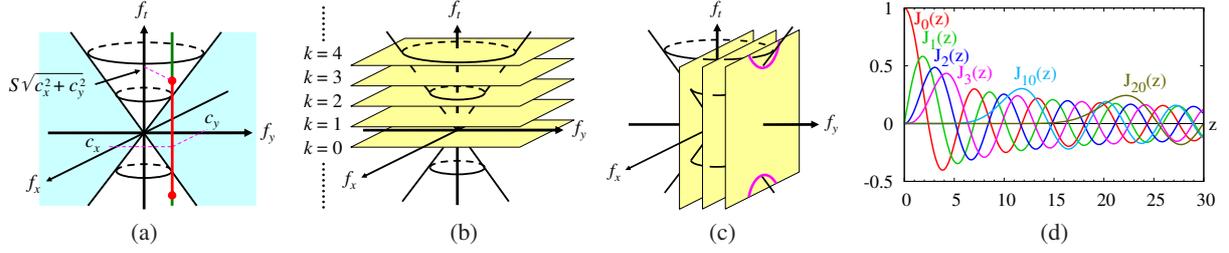


Figure 5: (a) The cone defining the volume (shown in blue) whose slices passing through the origin correspond to the power spectra of motion blur PSFs below the speed S . (b) Discrete f_i slices. (c) f_y slices. The hyperbolic intersections with the cone are shown in purple. (d) Plots of Bessel functions $J_k(z)$ of the first kind for some k , which correspond to the slices in (b).

where $J_k(z)$ is the k -th order Bessel function of the first kind [Wat22], which is plotted for some k in Fig. 5(d).

We show the *effectiveness* and *uniformity* of this distribution as described in Sec. 4.1. For effectiveness, we show $|\hat{p}(\mathbf{f}, f_i)|^2$ is small inside the cone $|f_i| \geq S|\mathbf{f}|$, shown in white in Fig. 5(a). By simple algebraic manipulation, we have $2\pi R|\mathbf{f}| < k$ inside the cone. As can be observed in Fig. 5(d) particularly clearly for $k = 10$ and 20 , Bessel functions $J_k(z)$ start from zero at the origin (except for $k = 0$), and remain small until coming close to the first maximum value, which is known to be around $z = k + 0.808618k^{1/3} > k$ [Wat22]. Therefore, $J_k(z)$ is small for $z < k$, which means $|\hat{p}(\mathbf{f}, f_i)|^2$ is small inside the cone.

Next, we show the uniformity of the distribution. Eq. (5) can be approximated using the asymptotic form of Bessel functions [Wat22] for $z \gg k^2$ as:

$$|\hat{p}(\mathbf{f}, f_i)|^2 \approx \frac{4}{\pi} \frac{T}{S|\mathbf{f}|} \cos^2 \left(2\pi R|\mathbf{f}| - \frac{k\pi}{2} - \frac{\pi}{4} \right). \quad (6)$$

This equation indicates that, at any given spatial frequency \mathbf{f} which is sufficiently large, $|\hat{p}(\mathbf{f}, f_i)|^2$ is a sinusoidal wave with an amplitude of $(4/\pi)(T/S|\mathbf{f}|)$, which is independent of f_i and hence uniform along the f_i direction. The amplitude itself is greater than the optimal assignment $T/S|\mathbf{f}|$ as described in Sec. 4.1, and averaging the cosine undulation in Eq. (6) reveals that the assigned frequency power is $(2/\pi)(T/S|\mathbf{f}|)$ on an average, meaning that the circular sensor motion achieves $2/\pi$ (about 64%) of the optimal assignment.

To verify the above argument, we show a numerically computed power spectrum of a spiral in Fig. 6 by three f_y slices as shown in Fig. 5(c), along with the power spectra of the other camera paths. The motion-invariant camera nearly optimally assigns the budget for the $f_y = 0$ slice corresponding to horizontal object motion, but it fails to deliver the budget uniformly for other cases. Our circular motion camera distributes the budget mostly evenly within the volume of interest, with condensed power around the cone surface corresponding to the maximum value of Bessel functions, which results in a moderate tendency to favor the target speed.

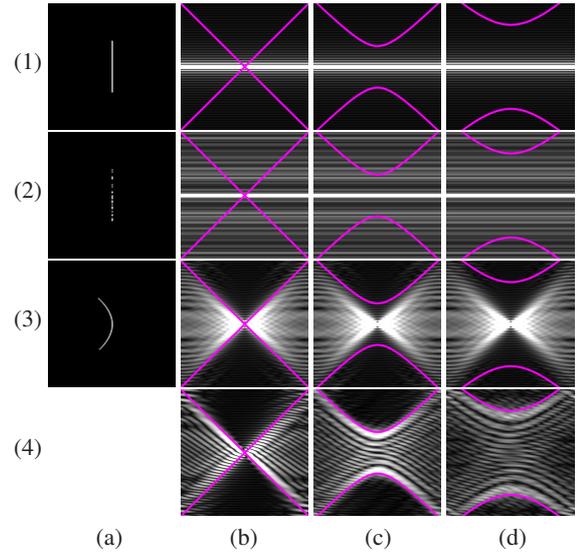


Figure 6: Camera paths in the space-time and 2D slices of their 3D log power spectra. Purple curves show the intersections with the cone of target speed S . Rows: (1) Static camera. (2) Coded exposure camera. (3) Motion-invariant camera. (4) Circular motion camera. Columns: (a) Camera path in the xt space-time. See Fig. 1(b) for the circular sensor motion path. (b) Slice at $f_y = 0$. (c)(d) Slices off the $f_x f_i$ plane ($f_y \neq 0$).

5. PSF Estimation

As shown in the bottom row of Fig. 2, the power spectra of PSFs resulting from circular sensor motion have different frequency zeros depending on object motion, serving as cues for PSF estimation. According to the model presented in [LFD07], PSF discriminability between candidate PSFs i and j can be measured using the following equation.

$$D(i, j) = \frac{1}{N} \sum_{f_x, f_y} \left(\frac{\sigma_i(f_x, f_y)}{\sigma_j(f_x, f_y)} - \log \frac{\sigma_i(f_x, f_y)}{\sigma_j(f_x, f_y)} \right) - 1, \quad (7)$$

where N is the number of discretized frequency components, and $\sigma_i(f_x, f_y)$ is the variance of the frequency component at (f_x, f_y) in images blurred with PSF i , which is given as:

$$\sigma_i(f_x, f_y) = \frac{\beta |F_i(f_x, f_y)|^2}{|G_x(f_x, f_y)|^2 + |G_y(f_x, f_y)|^2} + \eta, \quad (8)$$

where F_i and (G_x, G_y) are the Fourier transforms of PSF i and of the gradient operators, β is the variance of natural image gradients, and η is the noise variance (we set $\beta = 5.6 \times 10^{-3}$ and $\eta = 1.0 \times 10^{-6}$). $D(i, j)$ becomes large when the ratio of σ_i and σ_j is large, especially when either of them is zero (i.e., their frequency zero patterns are different).

To compare the PSF discriminability for various capture strategies, we generated a set of PSFs corresponding to all possible (discretized) object motions, and plot in Fig. 7(a) the minimum value of Eq. (7) among all the pairs of the PSFs. We set the target object speed as $S = 50$ pixel/sec, and considered object speed up to $1.5S$. Motion direction and speed were discretized by 15° and 5 pixels/sec, respectively. As shown by the red line, all of the capture strategies except ours have (almost) zero discriminability. This is because objects moving in opposite directions at the same speed produce (almost) the same motion blur except for the circular motion camera (see Fig. 7(b)(c)). We also plot the PSF discriminability (green line) apart from this 180° ambiguity by limiting object motion direction to $[0^\circ, 165^\circ]$. In this case, too, the circular motion camera gained the highest value.

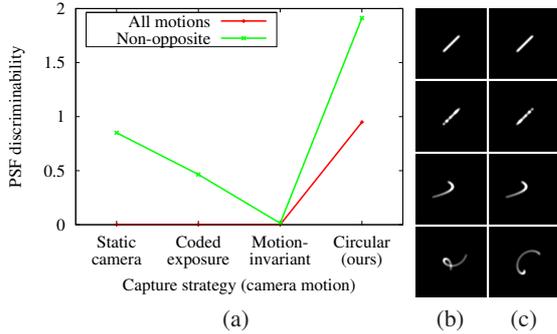


Figure 7: (a) Plot of PSF discriminability. (b) PSFs of 45° object motion direction for static, coded exposure, motion-invariant, and circular motion cameras (from top to bottom). (c) PSFs of 225° direction at the same speed.

Thanks to this high PSF discriminability, simple hypothesis testing works well in estimating PSFs for the circular motion camera, for which we used so-called MAP_k estimation [LWDF09]. We examine all possible object motions and pick the motion (equivalently the PSF) that gives the largest value for the following log posterior probability distribution.

$$\log p(F_i|B) = \sum_{f_x, f_y} \left[\log(\sigma_i(f_x, f_y)) + \frac{|B(f_x, f_y)|^2}{\sigma_i(f_x, f_y)} \right], \quad (9)$$

where B is the Fourier transform of a motion blurred image.

6. Experiments

Simulation: We evaluated the frequency preservation gained from the various image capture strategies by simulating motion blur for a set of 12 natural images, and by measuring the mean squared errors (MSE) between the Wiener-deconvolved images and the original unblurred images. Example images are shown in Fig. 8. Fig. 9 plots the deconvolution noise increase in decibels as $10 \log_{10}(MSE/\sigma^2)$, where

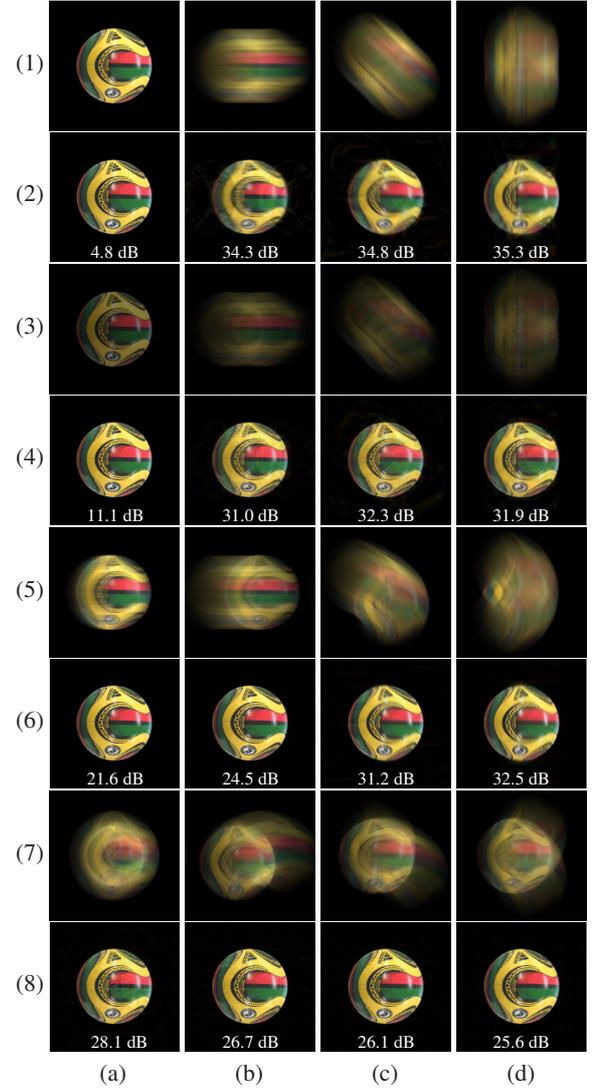


Figure 8: Simulated motion blurred images and their Wiener deconvolution results. The values indicate deconvolution noise increase. Rows: (1) Blurred and (2) deblurred images for a static camera. (3)(4) Coded exposure camera. (5)(6) Motion-invariant camera. (7)(8) Circular motion camera. Columns: (a) Static object. (b)(c)(d) Horizontal, oblique, and vertical object motion at the target speed S .

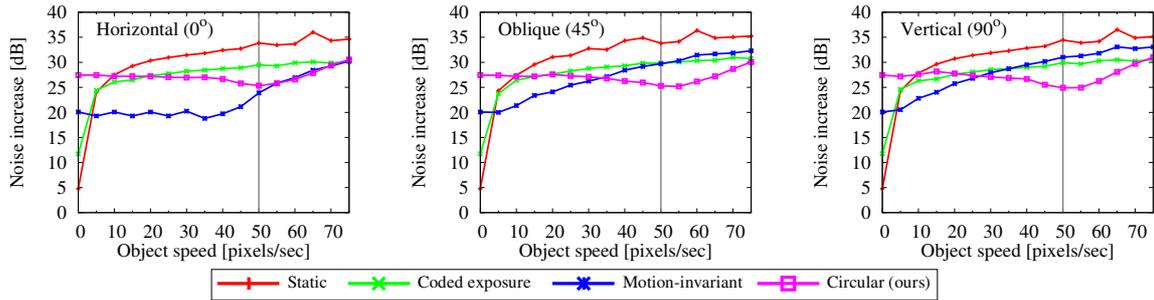


Figure 9: Plots of deconvolution noise increase for different object speeds and directions. The exposure time is 1 sec for all the cameras. The vertical gray lines indicate the target object speed $S = 50$ pixels/sec for the motion-invariant camera and ours. The length 50 code containing 25 ‘1’s [AXRT10] was used for the coded exposure camera (half the light level).

we assumed noise corruption for motion blur to be Gaussian of standard deviation $\sigma = 10^{-3}$ for $[0, 1]$ pixel values. The motion-invariant camera shows excellent constant performance for horizontal motion up to the target speed S , but for other directions, deconvolution noise increases for faster object motion. The coded exposure camera and ours have no such directional dependence. The coded exposure camera performs almost as perfectly as a static camera for static objects with marginal increase in deconvolution noise due to light loss, and the noise gradually increases for faster object motion. The circular motion camera also maintains stable performance up to and slightly beyond S . It moderately favors the target speed S , for which it has lower deconvolution noise than the other cameras except for the motion-invariant camera for horizontal object motion. The downside of our strategy is the increased noise for static objects.

Real examples using a prototype camera: For prototyping we placed a tilted acrylic plate inside the camera lens mount as shown in Fig. 10, and rotated it so that the refracted light rays were translated circularly. The plate is 3mm thick with a refraction index of 1.49, and the tilt angle is 7.7° , resulting in a circular motion radius R of 0.13mm. This corresponds to 5 pixels in our setup, and the target object speed is $S = 31.4$ pixels/sec with the exposure time $2T = 1.0$ sec.



Figure 10: Prototype camera based on a Canon EOS 40D. The lens is detached to reveal the modified lens mount. After passing through the lens, incoming light (shown in red) is displaced via the tilted acrylic plate, and the displacement sweeps a circle on the sensor while the plate rotates (yellow).

For deblurring, we performed the PSF estimation described in Sec. 5 for each user-segmented object, and applied the deconvolution method of Shan *et al.* [SJA08]. The deblurred objects and the background are blended back together. The PSF estimation took 20 min for a 512×512 image on an Intel Pentium 4 3.2GHz CPU.

Fig. 11 shows an example of multiple objects moving in different directions and at speeds. The digits and marks on the cars are visible in the deblurred image. For comparison, Fig. 12 shows closeups of the results from the static and circular motion camera images, in which we used simpler, Wiener deconvolution to better demonstrate high frequency preservation. More details were recovered for the circular motion camera image with less deconvolution noise.

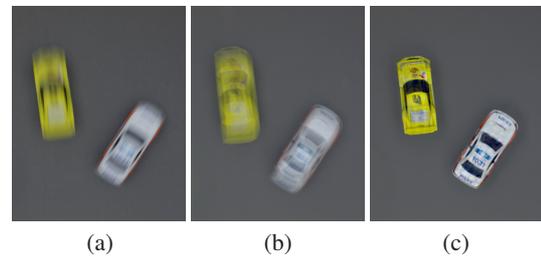


Figure 11: Toy cars. (a) From a static camera. (b) From the circular motion camera. (c) Deblurring result of (b).

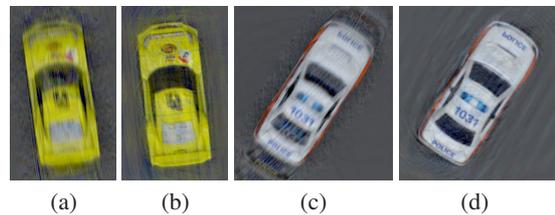


Figure 12: Comparison of Wiener deconvolution results for the toy car example. (a)(c) Results for the static camera image. (b)(d) Results for the circular motion camera image.

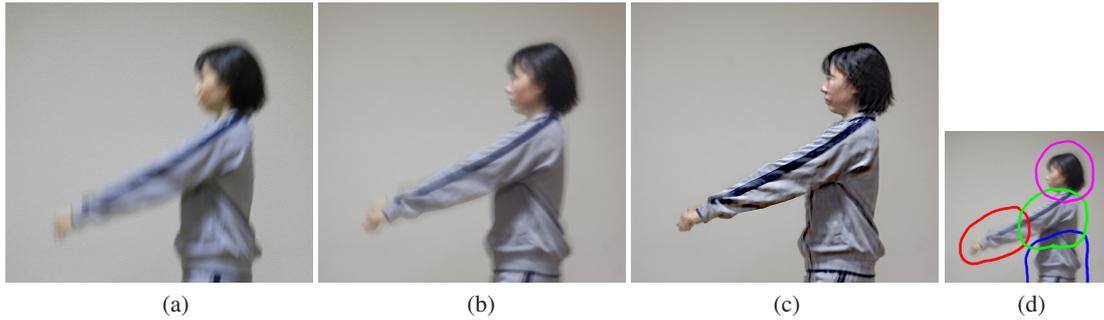


Figure 13: Squat motion. (a) From a static camera. (b) From the circular motion camera. (c) Deblurring result of (b). (d) User-specified motion segmentation. Four regions are enclosed by differently-colored lines.

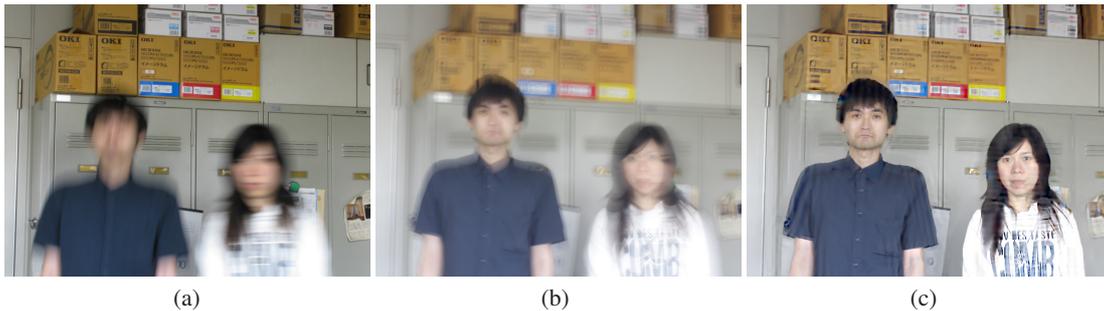


Figure 14: Moving people. (a) From a static camera. (b) From the circular motion camera. (c) Deblurring result of (b).

Fig. 13 shows an example of an object whose parts are moving differently. Fig. 13(d) shows the user-specified motion segmentation, for which manual specification took less than a minute. The regions overlap so that they can be stitched smoothly at the borders after deconvolution. Details such as fingers and wrinkles on the clothes were recovered.

Fig. 14 shows an example with a textured background. Due to occlusion boundaries, artifacts can be seen around the silhouettes of the people, but the deblurred faces are clearly recognizable. It is worth mentioning that the circular motion camera tells us that the man was moving downward while the woman was moving leftward (not upward or rightward), which is neither available information from the static camera image in Fig. 14(a) nor from the other capture strategies. We also note that details such as facial features are already visible in Fig. 14(b) even before deconvolution, which were successfully identified by an automatic facial feature point detector as shown in Fig. 15. These motion identification and recognizable image capture capabilities may be useful for surveillance purposes.

Comparison using a high-speed camera: For comparison with the other capture strategies, we used high-speed camera images of a horizontally moving resolution chart provided online [AXRT10]. Blurred images are simulated by averaging 150 frames from the 1,000 fps video, resulting in a 39-pixel blur. The length 50 code was used for the coded expo-

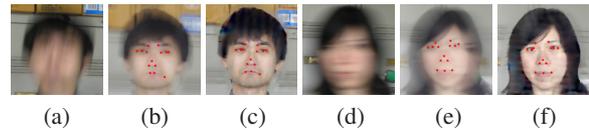


Figure 15: Results of facial feature point detection [YY08] for Fig. 14. (a)(d) Detection failed for the static camera image in Fig. 14(a), as the faces are severely blurred. (b)(e) Detection succeeded for the circular motion camera image in Fig. 14(b) even before they were deblurred, since the facial features are already visible. (c)(f) Detection also succeeded for the deblurred image in Fig. 14(c).

sure camera, spending 3 msec for each chop of the code. For fair comparison, the motion-invariant and circular motion cameras were targeted to an object speed of 50 pixels (not 39 pixels) per exposure time. We tilted the camera by 90° to simulate the “vertical” object motion relative to the camera. As shown in Fig. 16, the coded exposure deblurring produced a less noisy image than the static camera, but oblique streaks of noise can still be seen. The motion-invariant camera produced a clean image for horizontal object motion, but the result for vertical object motion exhibits severe noise. The circular motion camera produced clean images for both motion directions.

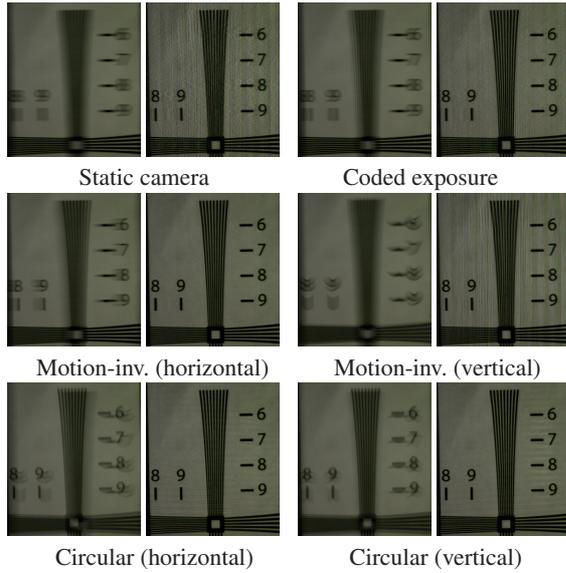


Figure 16: Comparison using high-speed camera images. For each pair of images, the left one is a simulated blurred image, and the right one is its deconvolution result. Please magnify the images on the PDF to clearly see the differences.

Subjective evaluation of recognizability: We have argued that motion-blurred objects are more recognizable with circular camera motion than the other image capture strategies. To quantitatively back up this claim, we conducted subjective evaluation where 55 persons were asked which of presented images were more recognizable to them (i.e., image textures and patterns such as facial features and text were more clearly seen or readable). We used the three images shown in Fig. 17, and we synthetically motion-blurred each image with the four image capture strategies as shown in Fig. 18. We presented every pair of the four blurred images to the subjects, and asked them to select either image (paired-comparison test). Therefore, 18 pairs were presented to the subjects (6 pairs for each image). This test was done four times with different object motion: static, horizontal, oblique, and vertical. From the number of “votes” from the subjects, relative recognizability can be quantified using Thurstone’s method [Thu27] as shown in Fig. 19. For static objects, recognizability was of course the best with a static camera, but for moving objects, the circular motion camera gained the highest values for all of the motion directions.



Figure 17: Original images used for the subjective evaluation, presented only once in the beginning of the test.



Figure 18: Blurred images used for the subjective evaluation. The shown images correspond to vertical object motion captured with static, coded exposure, motion-invariant, and circular motion cameras (from left to right). All of the six pairs of these four images were presented to the subjects.

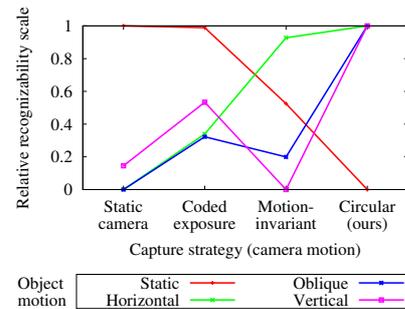


Figure 19: Relative recognizability scale between various image capture strategies.

7. Conclusions

We have proposed a method to facilitate motion blur removal from a single image by translating a camera sensor circularly about the optical axis during exposure, so that high frequencies can be preserved for a wide range of in-plane linear object motion within some target speed. We analyzed the frequency characteristics of circular sensor motion, and investigated its trade-offs between other image capture strategies. The advantages include reduced deconvolution noise at the target speed, improved PSF discriminability, and image recognizability without deconvolution.

The prototype implementation of the camera hardware may appear complicated, but it will be much simpler as technologies advance. In this paper we confined ourselves to in-plane linear object motion, and we also assumed user-specified motion segmentation. We would like to address these limitations in the future. Another issue of our method is that static objects are also blurred. One way to alleviate this is to pause the sensor for a fraction of exposure time. We intend to investigate ways to control the degree to which static and moving objects are favored relative to each other.

Acknowledgments

We thank Takeshi Naemura, Yusuke Iguchi, Yasutaka Furukawa, and the anonymous reviewers for their feedback; Paulo Silva, Pablo Garcia Trigo, Napaporn Metaaphanon, Yusuke Tsuda, and Saori Bando for their help; and those who kindly participated in the subjective evaluation.

References

- [AR09] AGRAWAL A., RASKAR R.: Optimal single image capture for motion deblurring. In *CVPR* (2009), pp. 2560–2567.
- [AX09] AGRAWAL A., XU Y.: Coded exposure deblurring: optimized codes for PSF estimation and invertibility. In *CVPR* (2009), pp. 2066–2073.
- [AXR09] AGRAWAL A., XU Y., RASKAR R.: Invertible motion blur in video. *ACM TOG* 28, 3 (2009), 95:1–95:8.
- [AXRT10] AGRAWAL A., XU Y., RASKAR R., TUMBLIN J.: Motion blur datasets and matlab codes. <http://www.umiacs.umd.edu/~aagrawal/MotionBlur/>, 2010.
- [BEN04] BEN-EZRA M., NAYAR S. K.: Motion-based motion deblurring. *IEEE Trans. PAMI* 26, 6 (2004), 689–698.
- [BEZN05] BEN-EZRA M., ZOMET A., NAYAR S. K.: Video super-resolution using controlled subpixel detector shifts. *IEEE Trans. PAMI* 27, 6 (2005), 977–987.
- [BLM90] BIEMOND J., LAGENDIJK R. L., MERSEREAU R. M.: Iterative methods for image deblurring. *Proceedings of the IEEE* 78, 5 (1990), 856–883.
- [Bra65] BRACEWELL R. N.: *The Fourier transform and its applications*. McGraw-Hill, 1965.
- [CL09] CHO S., LEE S.: Fast motion deblurring. *ACM TOG* 28, 5 (2009), 145:1–145:8.
- [CLDF10] CHO T. S., LEVIN A., DURAND F., FREEMAN W. T.: Motion blur removal with orthogonal parabolic exposures. In *IEEE Int. Conf. Computational Photo.* (2010), pp. 1–8.
- [DMY10] DING Y., MCCLOSKEY S., YU J.: Analysis of motion blur with a flutter shutter camera for non-linear motion. In *ECCV* (2010), pp. 15–30.
- [FSH*06] FERGUS R., SINGH B., HERTZMANN A., ROWEIS S. T., FREEMAN W. T.: Removing camera shake from a single photograph. *ACM TOG* 25, 3 (2006), 787–794.
- [Jia07] JIA J.: Single image motion deblurring using transparency. In *CVPR* (2007), pp. 1–8.
- [JKZS10] JOSHI N., KANG S. B., ZITNICK C. L., SZELISKI R.: Image deblurring using inertial measurement sensors. *ACM TOG* 29, 4 (2010), 30:1–30:8.
- [JSK08] JOSHI N., SZELISKI R., KRIEGMAN D.: PSF estimation using sharp edge prediction. In *CVPR* (2008), pp. 1–8.
- [KH96] KUNDUR D., HATZINAKOS D.: Blind image deconvolution. *IEEE Signal Processing Magazine* 13, 3 (1996), 43–64.
- [KTF11] KRISHNAN D., TAY T., FERGUS R.: Blind deconvolution using a normalized sparsity measure. In *CVPR* (2011), pp. 1–8.
- [Lev06] LEVIN A.: Blind motion deblurring using image statistics. In *Advances in Neural Information Processing Systems (NIPS)* (2006).
- [LDF07] LEVIN A., FERGUS R., DURAND F., FREEMAN W. T.: Image and depth from a conventional camera with a coded aperture. *ACM TOG* 26, 3 (2007), 70:1–70:9.
- [LSC*08] LEVIN A., SAND P., CHO T. S., DURAND F., FREEMAN W. T.: Motion-invariant photography. *ACM TOG* 27, 3 (2008), 71:1–71:9.
- [LWDF09] LEVIN A., WEISS Y., DURAND F., FREEMAN W. T.: Understanding and evaluating blind deconvolution algorithms. In *CVPR* (2009), pp. 1964–1971.
- [MLHR09] MOHAN A., LANMAN D., HIURA S., RASKAR R.: Image destablization: programmable defocus using lens and sensor motion. In *IEEE Int. Conf. Computational Photo.* (2009), pp. 1–8.
- [NKZN08] NAGAHARA H., KUTHIRUMMAL S., ZHOU C., NAYAR S. K.: Flexible depth of field photography. In *ECCV* (2008), pp. 60–73.
- [RAP05] RAV-ACHA A., PELEG S.: Two motion-blurred images are better than one. *Pattern Recog. Letters* 26, 3 (2005), 311–317.
- [RAT06] RASKAR R., AGRAWAL A., TUMBLIN J.: Coded exposure photography: motion deblurring using fluttered shutter. *ACM TOG* 25, 3 (2006), 795–804.
- [SJA08] SHAN Q., JIA J., AGARWALA A.: High-quality motion deblurring from a single image. *ACM TOG* 27, 3 (2008), 73:1–73:10.
- [SW71] STEIN E. M., WEISS G.: *Introduction to Fourier analysis on Euclidean spaces*. Princeton University Press, 1971.
- [SXJ07] SHAN Q., XIONG W., JIA J.: Rotational motion deblurring of a rigid object from a single image. In *ICCV* (2007), pp. 1–8.
- [TDBL08] TAI Y.-W., DU H., BROWN M. S., LIN S.: Image/video deblurring using a hybrid camera. In *CVPR* (2008), pp. 1–8.
- [Thu27] THURSTONE L. L.: A law of comparative judgement. *Psychological Review* 34, 4 (1927), 273–286.
- [TKLS10] TAI Y.-W., KONG N., LIN S., SHIN S. Y.: Coded exposure imaging for projective motion deblurring. In *CVPR* (2010), pp. 2408–2415.
- [Wat22] WATSON G. N.: *A treatise on the theory of Bessel functions*. Cambridge University Press, 1922.
- [XJ10] XU L., JIA J.: Two-phase kernel estimation for robust motion deblurring. In *ECCV* (2010), pp. 157–170.
- [YSQS07] YUAN L., SUN J., QUAN L., SHUM H.-Y.: Image deblurring with blurred/noisy image pairs. *ACM TOG* 26, 3 (2007), 1:1–1:10.
- [YSQS08] YUAN L., SUN J., QUAN L., SHUM H.-Y.: Progressive inter-scale and intra-scale non-blind image deconvolution. *ACM TOG* 27, 3 (2008), 74:1–74:10.
- [YY08] YUASA M., YAMAGUCHI O.: Real-time face blending by automatic facial feature point detection. In *IEEE Int. Conf. Automatic Face & Gesture Recognition* (2008), pp. 1–6.
- [ZGS10] ZHUO S., GUO D., SIM T.: Robust flash deblurring. In *CVPR* (2010), pp. 2440–2447.

Appendix A: Fourier Transform of a Spiral

According to [SW71], 2D Fourier transform of a function $g(r)e^{-ik\theta}$ is given as $G(f_r)e^{-ik\phi}$, where (r, θ) and (f_r, ϕ) are the polar coordinates in the primal and frequency domains, respectively (i.e., $f_r = |\mathbf{f}| \equiv |(f_x, f_y)|$), and we have:

$$G(f_r) = 2\pi i^{-k} \int_0^\infty g(r) J_k(2\pi f_r r) r dr. \quad (\text{A.1})$$

Applying this theorem to Eq. (4) with $k = 2\pi f_t / \omega$ leads to:

$$\begin{aligned} \hat{p}(\mathbf{f}, f_t) &= \left(2\pi i^{-k} \int_0^\infty \frac{\delta(r-R)}{R\omega} J_k(2\pi f_r r) r dr \right) e^{-ik\phi} \\ &= 2\pi i^{-k} \frac{1}{\omega} J_k(2\pi f_t R) e^{-ik\phi} \\ &= 2T J_k(2\pi R|\mathbf{f}|) i^{-k} e^{-ik\phi}. \end{aligned} \quad (\text{A.2})$$