Improved motion invariant imaging with time varying shutter functions

Steve Webster[a] and Andrew Dorrell[b]

Canon Information Systems Research, Australia (CiSRA),
Thomas Holt Drive, North Ryde, Australia

ABSTRACT

In motion invariant photography, blur is introduced by a structured movement of the camera during capture. The structured movement results in a uniform blur which simplifies deblur through post-processing for objects moving at different speeds in a single motion plane. However motion invariance depends on the camera speed exceeding the object speed (in the image plane) by a significant amount. This can lead to noisy image results and may be a problem for implementation of the method in practice. We propose the introduction of a time varying shutter transmittance to this recently proposed computational imaging method and demonstrate through simulation how this can improve both the degree of motion invariance and the reconstructed image quality, despite a reduction in optical efficiency. Improvements in the order 6dB are demonstrated for the reconstructed, deblurred images in the presence of moderate noise. The work has the potential to bring motion invariant photography closer to use in real camera product.

Keywords: Motion invariant, computational photography, motion deblur

1. INTRODUCTION

Motion blur occurs when the image focused onto the camera sensor changes during the exposure period, typically due to either camera shake or the motion of objects in the scene. Although motion blur can be used for artistic effect, it is generally regarded as a serious defect and results typically in images being discarded. While short exposure times are effective for reducing all forms of motion blur, they can result in increased noise and may not be practical if there is insufficient scene illumination. For longer exposure periods or scenes containing significant motion, blurring of the captured image will still occur. As a result, there has been long-term interest in post-processing to remove motion blur from images.[1,2]

For motion in a single orientation plane, Levin, Sand, Cho, Durand and Freeman[3] identified that additional, deliberate blurring of the image by motion of the camera relative to the scene could be used to achieve a substantially uniform blurring of both the scene and moving objects in the scene. This means that a single blur kernel can be used to deconvolve all motion blur in the captured image. The method has theoretical connections with wavefront coding,[4] which rather than motion blur, addresses blur due to defocus of the optical system. Analogously, it extends the depth of focus of an optical system by introducing a depth independent defocus blur. The optimality of the motion invariant approach when compared to other computational approaches to motion deblur such as the flutter shutter[5] has been considered elsewhere.[6]

In this paper we address the effect of real world constraints on motion invariant photography, specifically the speed and extent of camera motion for the given exposure time. We investigate the use of time varying shutter functions for improving the motion invariance of the method. In section 2 we briefly review the theory behind restoration of images with motion blur and motion invariant photography; in section 3 we describe our improvement which is to introduce a time varying shutter function to the motion invariant capture process and calculate the corresponding blur kernels. Kernel analysis and motion deblur simulation results are then presented in section 4 before a discussion in section 5 and conclusions in section 6. The results indicate that motion invariance is increased by smooth temporal windowing, and that improved conditioning permits the use of relatively simple deconvolution methods.

Further author information:
[a] Email: steve.webster@cisra.canon.com.au, Telephone: +61 (2) 9805 2245
[b] Email: andrew.dorrell@cisra.canon.com.au, Telephone: +61 (2) 9805 2224

Digital Photography VII, edited by Francisco H. Imai, Feng Xiao, Jeffrey M. DiCarlo,
Nitin Sampat, Sebastiano Battiato, Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 7876, 787604

2. BACKGROUND

It is common to model a blurred image as a convolution of an ideal image:

$$g(x, y) = f(x, y) \otimes h(x, y) + \eta(x, y) \tag{1}$$

where g is the blurred image, f is the blur-free image, h is the kernel or point spread function (PSF) used to model the motion blur and η is noise. Using this model, an estimate of f is recovered from g by inverting h. In the Fourier domain, direct inversion gives

$$F(u, v) + H^{-1}N(u, v) = H^{-1}G(u, v). \tag{2}$$

where capitalisation denotes Fourier space representation. If the blur kernel is known then restoration is relatively straightforward and provides images of useful visual quality while errors in the kernel estimate will result in structured artifacts such as ringing. In general however, estimation of the blur kernel is difficult. Blur typically varies across the image, being a function of the motion speed and direction. It becomes especially complex where objects are moving at different speeds in the scene. In such cases, iterative blind deconvolution methods[7] are required to achieve useful results. These methods are generally complex and time consuming.

Conditioning of the blur kernel also plays a significant role in determining the quality of the restored image. As H is typically low-pass, H^{-1} will amplify high frequency noise and contaminate the output. The condition number of the blur function is defined as

$$\kappa(H) = \|H\| \cdot \|H^{-1}\| \tag{3}$$

where $\|\cdot\|$ denotes the operator norm. Condition indicates, in the worst case, how much the function can change in proportion to small changes in the argument. It is indicative of the operator's sensitivity to noise. A larger condition number indicates higher sensitivity to noise. This may manifest, for example, as significant notches in an operator's Fourier spectrum — which would lead to significant attenuation of the signal towards the noise floor at certain frequencies.

Traditional deconvolution methods such as Wiener and Lucy-Richardson[8,9] include assumptions about the noise model that help to minimise sensitivity to noise. However a blur function that exhibits a relatively smooth spectral response over a wide frequency band will be better conditioned and should be a design goal for any computational system based on deconvolution.

When modelling capture noise for the evaluation of sensor independent computational imaging methods, it is usually sufficient to consider a simplified model comprising two terms: data dependent (photon shot) noise and data independent (sensor read) noise. The latter is modelled as zero mean Gaussian for simplicity. The noise model relates the measured image, $g(x, y)$, to an ideal (continuous in [0,1] and noise free) linear intensity image, $f(x, y)$, of the scene according to

$$g(x, y) = \frac{\text{Poisson}(pf(x, y))}{p} + \text{Gaussian}\left(0, \left(\frac{\sigma}{p}\right)^2\right) \tag{4}$$

where p is the maximum number of photo-electrons for any pixel and σ is the standard deviation corresponding to the read noise. Scaling σ by $1/p$ ensures that the significance of sensor noise increases as the photon catch decreases. The model is similar that used in other recent computational imaging analyses.[6]

2.1 Motion invariant photography

The "motion invariant photography" method[3] generates a uniform motion blur for all objects undergoing translation in a single orientation plane in the scene. To achieve this, additional blur is introduced by constant 1D acceleration of the camera, parallel to the plane of scene motion, during exposure such that the plot of camera position with time follows a parabola. Although prototyped using a moving camera, Levin et al.[3] note that a translation of the optical image relative to the sensor can be achieved using various image stabilisation technologies such as shifting of optical elements and shifting of the sensor.[10] Due to the constant acceleration of

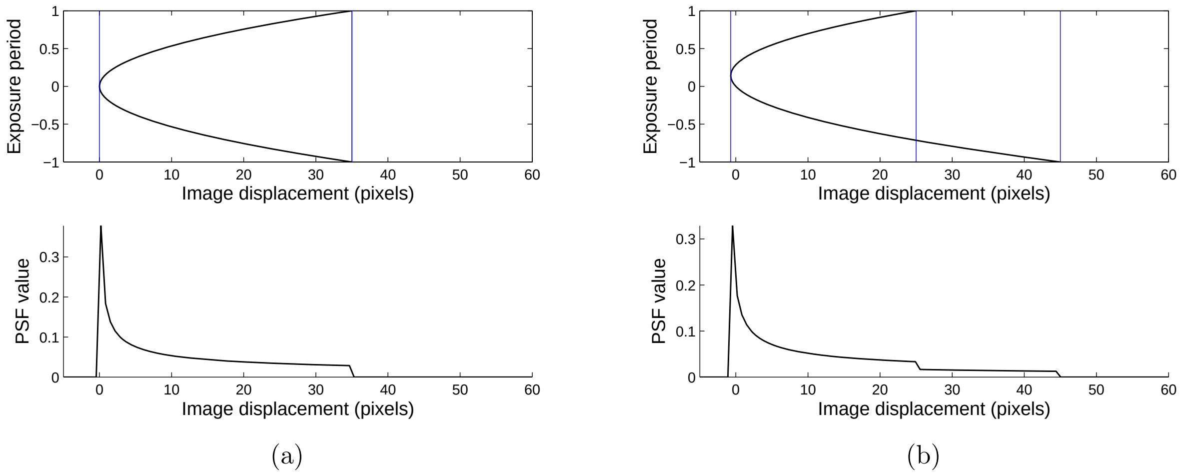


Figure 1. Shuttering in the camera imposes an envelope over the integration which affects the resulting blur kernel (PSF). For a conventional shutter, abrupt discontinuities result in sharp differences between the PSF corresponding to stationary (a) and moving (b) objects.

the camera between an equal and opposite start and finish speed, each moving object is tracked equally by the camera for a brief time during exposure, ensuring each is blurred equally. Formally, the offset x as a function of time t of a point source, moving relative to the scene with speed s on a sensor undergoing parabolic translation with acceleration a is given by

$$x(t) = at^2 - st \tag{5}$$

or after completing the square, by

$$x(t) = a\left(t - \frac{s}{2a}\right)^2 - \frac{s^2}{4a} \tag{6}$$

which corresponds to the family of parabolas with a time offset of $-s/2a$ and spatial offset of $-s^2/4a$. The blur kernel that results from this motion on the sensor can be obtained from the inverse curve,

$$t(x) = \frac{s \pm \sqrt{s^2 + 4ax}}{2a}, \tag{7}$$

by calculating its slope. This corresponds to the "amount of time" the camera spent at each offset x. Thus, for an unbounded parabolic motion of the camera

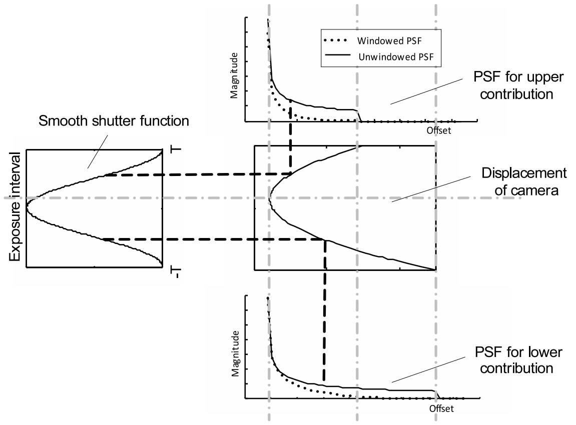
$$h(x) = \frac{1}{\sqrt{a\left(x + \frac{s^2}{4a}\right)}}. \tag{8}$$

The family of curves $h(x)$ are all translations of the curve $1/\sqrt{ax}$ so the blur kernel is invariant up to a shift of $-s^2/4a$

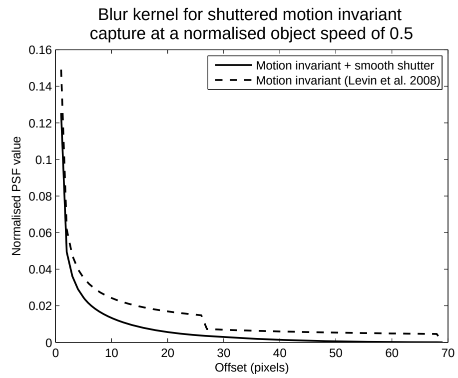
In practice the parabola is bounded over a fixed exposure period of $[-T, T]$. In this case, discontinuities arise as shown in Figure 1 and the blur kernel is defined piecewise as

$$h(x) = \begin{cases} \frac{1}{\sqrt{a\left(x+\frac{s^2}{4a}\right)}} & \text{if } -\frac{s^2}{4a} \leq x < aT^2 - sT \\ \frac{1}{2\sqrt{a\left(x+\frac{s^2}{4a}\right)}} & \text{if } aT^2 - sT \leq x \leq aT^2 + sT \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

From the above we see that there will be some motion dependence in the blur kernel and the degree of this motion dependence is determined by the truncation of the parabolic path due to the limited exposure window. If



(a)



(b)

Figure 2. Calculation of the shuttered blur kernel involves using a window function to weight the gradient of the inverse curve(a). The resulting kernels (b) show differences in the shape of the tail. In particular, smooth window functions signficantly reduce the discontinuities at the cost of some overall signal attenuation.

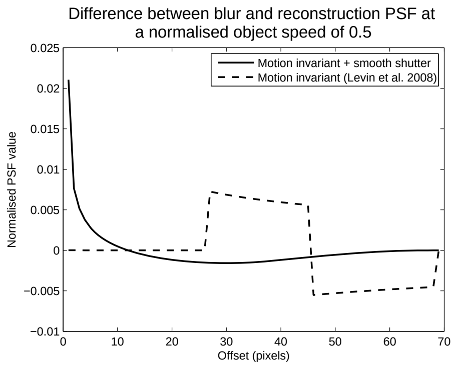
we use a single kernel for deconvolution, there will be a motion dependent error in the image reconstruction. It is possible to minimize the impact of these discontinuities by extending shutter period, or equivalently increasing the acceleration a. When the range of translation speeds achieved by the camera exceeds the range of speeds in the scene by a sufficient amount, then the discontinuities introduced by shuttering will appear further towards the tail of the blur kernel and their impact will be less significant. While acknowledging these discontinuities, Levin et al.[3] did not analyse their effect in detail. In practical implementations of a motion invariant camera, these discontinuities may become significant. In particular, it may not be possible to achieve the range of translation speeds required for good invariance within an appropriate exposure time without excessive displacement of the camera. In addition, increasing the acceleration without increasing the exposure time means that each part of the scene is "tracked" for a shorter time leading to a decreased signal to noise ratio in the captured image and leading to excessive noise in the reconstructed (deblurred) images.

3. IMPROVING MOTION INVARIANT CAPTURE

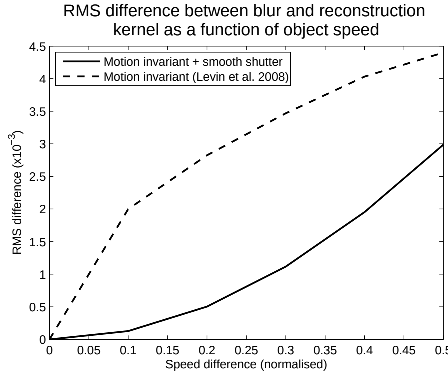
To improve the performance of motion invariant photography we begin by observing that conventional shuttering in a camera can be modelled as a 'rect' function. In place of this conventional shutter we propose using a time varying transmittance (shutter) function. The shutter function should implement a smooth temporal window. The use of a smooth windowing function has the advantage of reducing the discontinuity in the PSF that results from shuttering. By selecting the shutter function appropriately we are also able to optimise the Fourier domain properties of the blur kernel. In this paper we consider functions borrowed from the field of digital filter design.

The shuttered blur kernel was numerically evaluated in two parts corresponding to the upper and lower portions of the parabolic path as depicted in Figure 2. For each sample point in the kernel the upper and lower contributions are calculated by area integration and then weighted according to the corresponding value of the shutter function. The blur kernel sample is determined as the sum of these contributions. The coefficients for the smoothly shuttered kernel are then normalised according to the sum of coefficients in the equivalent 'rect' shuttered kernel in order to account for the loss of optical efficiency.

In our simulation, object speed at the image plane is defined relative to a stationary image sensor and normalised to the maximum speed of the camera / sensor. The blur kernels corresponding to a range of object speeds up to half of the maximum camera speed were evaluated. These ranged in length from 34-51 samples. The relative displacement between the input and reconstructed image, as predicted by equation (6) was removed from the calculated kernel for simplicity and to permit direct comparison of input and output images for PSNR evaluation. The assumed blur kernel, used for restoration, is the kernel corresponding to a stationary object — i.e. one with zero speed. Our study focuses on the differences between the effective blur kernel for a non-stationary subject and the assumed blur kernel along with the impact of those differences on image restoration.

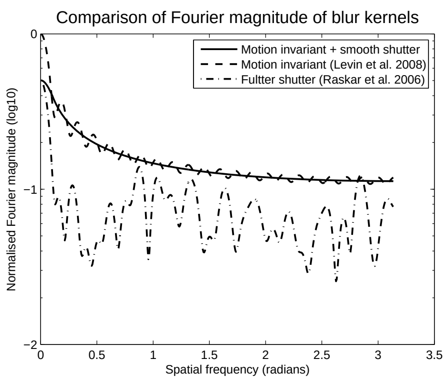


(a)



(b)

Figure 3. Difference between the blur kernel for an object moving at a normalised speed of 0.5 and a stationary object is depicted for the 'rect' shutter and the smooth Hann shutter (a). The RMS difference as a function of speed is graphed in (b). Speeds are defined as the speed of the object's image at the image plane relative to a stationary image sensor, normalised to the maximum speed of the camera / sensor.



(a)

Blur kernel	Condition
Motion invariant + smooth shutter	4.45
Motion invariant (Levin et al. 2008)	9.27
Flutter shutter (Raskar et al. 2006)	19.5

(b)

Figure 4. Fourier magnitude (a) and condition number (b) of various blur kernels.

4. RESULTS

4.1 Kernel analysis

Figure 3 shows that the differences between the effective blur kernel for a moving object and the assumed blur kernel used for deconvolution are localised near the origin and drop quickly away when smooth shuttering is used. By comparison, the conventional shutter approach produces errors at a distance. In practice this results in "ghosting" in the reconstructed images. The RMS differences between the blur kernel for a moving object and a stationary object are graphed in Figure 3(b). For 'rect' shuttering of the motion invariant capture, a steady increase in kernel difference with object speed is observed. Smooth shuttering both reduces the rate at which this difference grows and significantly reduces the difference for a range of object speeds. Note that, as the range of speeds is defined relative to the notional speed of the camera system during capture, these results generalise over the range of parameters considered.

Figure 4(a) shows the magnitude of the Fourier spectrum of the motion invariant blur kernel with 'rect' shuttering along with the spectrum of the smoothly shuttered blur kernel proposed in this paper. For reference, we have also shown the spectrum of the blur kernel arising from the "flutter shutter" function proposed by Raskar[5] (which is not motion invariant) for a normalised speed of 0.5. In general the motion invariant method has a better frequency characteristic, having less attenuation and a relatively flat response compared to this previous, spatially variant method. The introduction of a smooth shutter makes the response of the motion invariant method notably flatter, however there is attenuation in the low frequencies. This is a direct consequence of smooth shuttering, which reduces the optical efficiency of capture. Note that the attenuation in the lower frequencies means the overall response of the blur kernel for smoothly shuttered motion invariant capture shows much less variation. Correspondingly, numerical evaluation of the condition numbers for the blur kernels, tabulated in Figure 4, indicates a potential reconstruction benefit for smooth shuttering.

We note that, in general, the conditioning of the motion invariant system is notably better than previously proposed binary modulation functions and the introduction of smooth shuttering reduces the condition number by a further factor of two. As condition number is an indication only of worst case noise gain however, it is instructive to consider a simulation of capture and restoration applied to image data.

4.2 Capture simulation

To simulate image capture by a camera, input images were first converted to an approximation of linear intensity by inverting the gamma coding (a standard value of 2.2 is assumed in this step) prior to convolving with the numerically evaluated blur kernel. The blurred images were then contaminated with noise to provide pixel data analogous to what would be read from a camera sensor prior to deconvolution and gamma (re-)coding for visualization and PSNR evaluation. Noise was applied according to the model of equation (4). Deconvolution was performed using the standard Matlab implementations of Wiener, Richardson-Lucy and regularised deconvolution. In all cases the blur kernel corresponding to an object speed of zero was used as the kernel argument for deconvolution.

Simulation was performed for a standard 'rect' shutter function as well as a number of smooth window functions including triangle, Hann and Gaussian windows for a range of simulated object speeds and noise levels. In each case a constant exposure time is assumed. As noted above, the smoothly shuttered kernels were normalised for the total exposure when using a 'rect' shutter so that the sacrifice in optical efficiency that results from temporal windowing is reflected in the simulated raw data. In practice this means that the effect of noise contamination is greater for the smoothly shuttered source data. This gives a more demanding baseline for comparison.

The results of a simulated capture and reconstruction by 30 iterations of Richardson-Lucy deconvolution are shown in Figure 5 for the well known "bird" image. We observe that the discontinuities in the conventional 'rect' shuttered capture lead to ghosting artifacts in the reconstructed image as shown in Figure 5(b–c). This ghosting was found to be present with other deconvolution algorithms including the Wiener and regularized deconvolution methods also provided in Matlab. The results for smoothly shuttered capture, as depicted in Figure 5(e–f) for the Hann shutter function, show no sign of ghosting artifact.

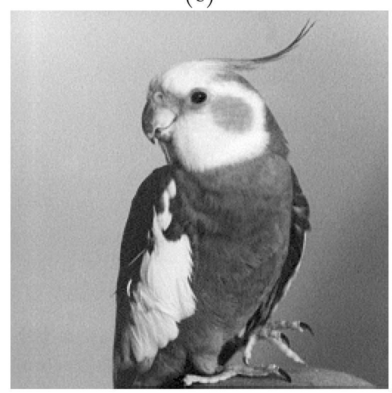
To determine the relative severity of reconstruction errors, the PSNR was determined for a range of (simulated) object motion speeds. Calculations were made for a range of noise levels and window functions including the conventional 'rect' function. Selected PSNR results for the bird image are shown in Figure 6. The shapes of the plots are representative of the results across the range of parameter values considered. The results indicate that smooth shuttering provides a useful advantage for almost all speeds and noise levels. The most significant improvements are observed in the mid to high object speeds where they are typically in the order of 6dB. The Hann window function yielded the best PSNR results of the shutter functions considered. The PSNR differences for different smooth window functions were small (< 1dB) when compared to the 3–8dB difference when compared to 'rect' shuttering. The results exemplified in Figure 6(a) show that smooth shuttering leads to a relatively consistent reconstruction PSNR for speeds up to around 0.2 after which performance declines. By comparison, the performance of 'rect' shuttering declines approximately in proportion to speed. This is consistent with improved motion invariance for the smooth shuttered method. The performance of the system with respect to noise as exemplified in Figure 6(b) is also improved by smooth shuttering. PSNR drops sharply as the signal approaches the noise floor irrespective of the shuttering used, though this drop off is faster for smooth shuttering due to the reduced exposure level.

For smoothly shuttered motion invariant capture, PSNR is improved for all object speeds, including zero (stationary). This would appear to result from the improved conditioning of the blur kernel which makes it less sensitive to noise in the input image.



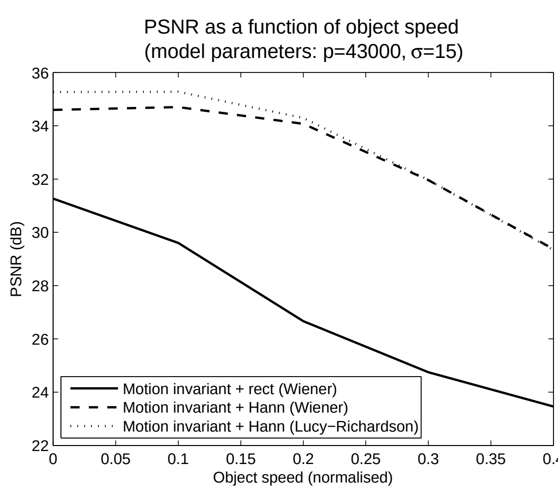
(a) (b) (c)



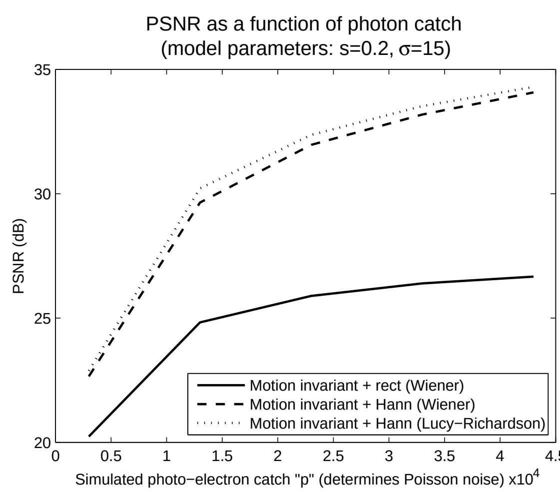


(d) (e) (f)

Figure 5. Simulation results using the bird image(a). The image (d) shows overlayed the result of simulated motion invariant capture using a rect shutter (top) and Hann shutter (bottom) for comparison. The loss of optical efficiency results in the Hann shuttered image being darker. The remaining images provide a comparison of deblurred output for the cases of a conventional rect shutter at relative speeds of 0.1 (b) and 0.3 (c) with the cases of a Hann shutter for the same speeds (e–f).



(a)



(b)

Figure 6. PSNR results from simulation of capture and restoration as a function of (a) object speed and (b) noise.

5. DISCUSSION

In a practical implementation of a motion invariant camera it is desirable to limit the range of camera translation. This is both because of practical optical and mechanical issues and because a smaller range of translation will result in better use of the available sensor area. It has been suggested that parabolic translation in time of the camera relative to the scene may be implemented using movable lens element or that the sensor itself could be translated. Both these schemes would seem plausible given their current use in image stabilization.[10] In practice these schemes can only translate the image by modest amounts, making the smoothly shuttered method proposed herein more important. Smooth shuttering itself may be implemented using a variety of electronic and mechanical means. Mechanical means such as an iris may complicate the restoration problem due to their effect on the PSF of the optical system. For this reason something like an LCD shutter will typically provide a more desirable implementation.

6. CONCLUSION

We have considered a solution to some of the practical issues with the motion invariant imaging method proposed by Levin et al. In particular, we have considered practical constraints on the exposure time as well as speed and extent of the excursion permitted at the camera. These result in motion dependent variations in the blur kernel that lead to ghosting artifacts, especially when using standard deconvolution methods. We have shown that smooth shutter functions can significantly reduce this artifact and also lead to better conditioned deconvolution problems and improvements to reconstruction PSNR. The results of both kernel analysis and camera simulation are also consistent with smooth shuttering providing increased motion invariance. The results should make the motion invariant imaging method more interesting for use in practical camera systems.

REFERENCES

[1] Kay, I. W. and Cochran, G. D., "Device for removing the effect of blurring in photography," U.S. Patent 3512464, Conductron Corporation (May 1970). (Filed Nov 21, 1966).

[2] Shan, Q., Jia, J., and Agarwala, A., "High-quality motion deblurring from a single image," *ACM Trans. Graph.* **27**(3), 1–10 (2008).

[3] Levin, A., Sand, P., Cho, T. S., Durand, F., and Freeman, W. T., "Motion-invariant photography," in [*ACM SIGGRAPH 2008 papers*], 1–9, ACM, Los Angeles, California (2008).

[4] Dowski, J. and Cathey, W. T., "Extended depth of field through wave-front coding," *Applied Optics* **34**, 1859–1866 (Apr. 1995).

[5] Raskar, R., Agrawal, A., and Tumblin, J., "Coded exposure photography: motion deblurring using fluttered shutter," in [*ACM SIGGRAPH 2006 Papers*], *SIGGRAPH '06*, 795–804, ACM, New York, NY, USA (2006).

[6] Agrawal, A. and Raskar, R., "Optimal single image capture for motion deblurring," in [*Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*], 2560–2567 (2009).

[7] Levin, A., Weiss, Y., Durand, F., and Freeman, W., "Understanding and evaluating blind deconvolution algorithms," in [*Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*], 1964–1971 (2009).

[8] Richardson, W. H., "Bayesian-Based iterative method of image restoration," *Journal of the Optical Society of America* **62**, 55–59 (Jan. 1972).

[9] Lucy, L. B., "An iterative technique for the rectification of observed distributions," *Astronomical Journal* **79**, 745 (June 1974).

[10] Levoy, M., "Optical image stabilization (IS)," course notes CS 448A, Computer Science Department Stanford University (2010).