Flexible Depth of Field Photography

Sujit Kuthirummal, Hajime Nagahara, Changyin Zhou, and Shree K. Nayar

Abstract—The range of scene depths that appear focused in an image is known as the depth of field (DOF). Conventional cameras are limited by a fundamental trade-off between depth of field and signal-to-noise ratio (SNR). For a dark scene, the aperture of the lens must be opened up to maintain SNR, which causes the DOF to reduce. Also, today's cameras have DOFs that correspond to a single slab that is perpendicular to the optical axis. In this paper, we present an imaging system that enables one to control the DOF in new and powerful ways. Our approach is to vary the position and/or orientation of the image detector, *during* the integration time of a single photograph. Even when the detector motion is very small (tens of microns), a large range of scene depths (several meters) is captured both in and out of focus.

Our prototype camera uses a micro-actuator to translate the detector along the optical axis during image integration. Using this device, we demonstrate four applications of flexible DOF. First, we describe extended DOF, where a large depth range is captured with a very wide aperture (low noise) but with nearly depth-independent defocus blur. Deconvolving a captured image with a single blur kernel gives an image with extended DOF and high SNR. Next, we show the capture of images with discontinuous DOFs. For instance, near and far objects can be imaged with sharpness while objects in between are severely blurred. Third, we show that our camera can capture images with tilted DOFs (Scheimpflug imaging) without tilting the image detector. Finally, we demonstrate how our camera can be used to realize non-planar DOFs. We believe flexible DOF imaging can open a new creative dimension in photography and lead to new capabilities in scientific imaging, vision, and graphics.

Index Terms—I.4.1.b Imaging geometry, programmable depth of field, detector motion, depth-independent defocus blur

◆

1 DEPTH OF FIELD

The depth of field (DOF) of an imaging system is the range of scene depths that appear focused in an image. In virtually all applications of imaging, ranging from consumer photography to optical microscopy, it is desirable to control the DOF. Of particular interest is the ability to capture scenes with very large DOFs. DOF can be increased by making the aperture smaller. However, this reduces the amount of light received by the detector, resulting in greater image noise (lower SNR). This trade-off gets worse with increase in spatial resolution (decrease in pixel size). As pixels get smaller, DOF decreases since the defocus blur occupies a greater number of pixels. At the same time, each pixel receives less light and hence SNR falls as well. This trade-off between DOF and SNR is one of the fundamental, long-standing limitations of imaging.

In a conventional camera, for any location of the image detector, there is one scene plane – the focal plane – that is perfectly focused. In this paper, we propose varying the position and/or orientation of the image detector *during* the integration time of a photograph. As a result, the focal plane is swept through a volume of the scene causing all points within it to come into and go out of focus, while the detector collects photons.

We demonstrate that such an imaging system enables one to control the DOF in new and powerful ways:

- **Extended Depth of Field:** Consider the case where a detector with a global shutter (all pixels are exposed simultaneously and for the same duration) is moved with *uniform speed* during image integration. Then, each scene point is captured under a continuous range of focus settings, including perfect focus. We analyze the resulting defocus blur kernel and show that it is nearly constant over the range of depths that the focal plane sweeps through during detector motion. Consequently, the captured image can be deconvolved with a single, known blur kernel to recover an image with significantly greater DOF, without having to know or determine scene geometry. This approach is similar in spirit to Hausler's work in microscopy [1]. He showed that the DOF of an optical microscope can be enhanced by moving a specimen of depth range d, a distance $2d$ along the optical axis of the microscope, while filming the specimen. The defocus of the resulting captured image is similar over the entire depth range of the specimen. However, this approach of moving the scene with respect to the imaging system is practical only in microscopy and not suitable for general scenes. More importantly, Hausler's derivation assumes that defocus blur varies linearly with scene depth which is true only for imaging systems that are telecentric on the object side such as microscopes.
- **Discontinuous Depth of Field:** A conventional camera's DOF is a single fronto-parallel slab located around the focal plane. We show that by moving a global-shutter detector *non-uniformly*, we can capture images that are focused for certain specified scene depths, but defocused for in-between scene regions. Consider a scene that

---

- *Sujit Kuthirummal is with Sarnoff Corporation, Princeton, NJ. USA. 08540.*
- *Hajime Nagahara is with the Graduate School of Engineering Science, Osaka University, Osaka, Japan. 5608531.*
- *Changyin Zhou and Shree K. Nayar are with the Department of Computer Science, Columbia University, New York, NY. USA. 10027.*

includes a person in the foreground, a landscape in the background, and a dirty window in between the two. By focusing the detector on the nearby person for some duration and the far away landscape for the rest of the integration time, we get an image in which both appear fairly well-focused, while the dirty window is blurred out and hence optically erased.

- **Tilted Depth of Field:** Most cameras can only focus on a fronto-parallel plane. An exception is the view camera configuration [2], [3], where the image detector is tilted with respect to the lens. When this is done, the focal plane is tilted according to the well-known Scheimpflug condition [4]. We show that by *uniformly* translating an image detector with a rolling electronic shutter (different rows are exposed at different time intervals but for the same duration), we emulate a tilted image detector. As a result, we capture an image with a tilted focal plane and hence a tilted DOF.
- **Non-planar Depth of Field:** In traditional cameras, the focal surface is a plane. In some applications it might be useful to have a curved / non-planar scene surface in focus. We show that by *non-uniformly* (with varying speed) translating an image detector with a rolling shutter we emulate a non-planar image detector. As a result, we get a non-planar focal surface and hence a non-planar DOF.

An important feature of our approach is that the focal plane of the camera can be swept through a large range of scene depths with a very small translation of the image detector. For instance, with a 12.5 mm focal length lens, to sweep the focal plane from a distance of 450 mm from the lens to infinity, the detector has to be translated only about 360 microns. Since a detector only weighs a few milligrams, a variety of micro-actuators (solenoids, piezoelectric stacks, ultrasonic transducers, DC motors) can be used to move it over the required distance within very short integration times (less than a millisecond if required). Note that such micro-actuators are already used in most consumer cameras for focus and aperture control and for lens stabilization. We present several results that demonstrate the flexibility of our system to control DOF in unusual ways.

This is the extended version of a paper that appeared in [5].

2 RELATED WORK

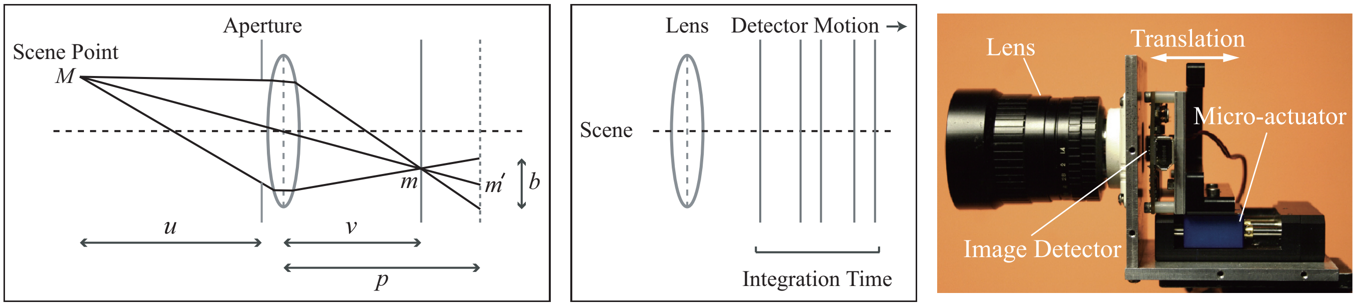
In microscopy, Hausler [1] demonstrated that DOF can be extended by changing the focus during image integration – by moving the specimen. We also propose changing the focus during image integration, but by moving the image detector. We show that for conventional imaging geometries, a particular detector motion – constant velocity – enables us to realize extended DOF. As mentioned above, Hausler's work assumes that defocus blur varies linearly with scene depth, which is true for imaging systems that are telecentric on the object side, like microscopes. On the other hand, our approach, for conventional (non-telecentric) imaging geometries, assumes that defocus blur varies linearly with the translation of the detector. Note that though the two approaches are for different imaging geometries, they make the same underlying assumption that defocus blur varies linearly with axial translation of a particular element – the scene (in Hausler's work) or the detector (in ours). While constant velocity detector motion enables extended DOF, we also show how other detector motions enable different DOF manipulations like discontinuous, tilted, and non-planar DOFs.

A promising approach to extended DOF imaging is wavefront coding, where phase plates placed at the aperture of the lens cause scene objects within a certain depth range to be defocused in the same way [6], [7], [8]. Thus, by deconvolving the captured image with a single blur kernel, one can obtain an all-focused image. The effective DOF is determined by the phase plate used and is fixed. On the other hand, in our system, the DOF can be chosen by controlling the motion of the detector.

Recently, Levin et al. [9] and Veeraraghavan et al. [10] have used masks at the lens aperture to control the properties of the defocus blur kernel. From a single captured image, they aim to estimate the structure of the scene and then use the corresponding depth-dependent blur kernels to deconvolve the image and get an all-focused image. However, they assume simple layered scenes and their depth recovery is not robust. In contrast, our approach is not geared towards depth recovery, but can significantly extend DOF. Also, the masks used in these previous works attenuate some of the light entering the lens, while our system operates with a clear and wide aperture. All-focused images can also be computed from an image captured using integral photography [11], [12], [13]. However, since these cameras make spatio-angular resolution trade-offs to capture 4D lightfields in a single image, the computed images have much lower spatial resolution when compared to our approach.

A related approach is to capture many images to form a focal stack [14], [15], [16]. An all-in-focus image as well as scene depth can be computed from a focal stack. However, the need to acquire multiple images increases the total capture time making the method suitable for only quasi-static scenes. An alternative is to use very small exposures for the individual images. However, in addition to the practical problems involved in reading out the many images quickly, this approach would result in under-exposed and noisy images that are unsuitable for depth recovery. Recently, Hasinoff and Kutulakos [17] have proposed a technique to minimize the total capture time of a focal stack, given a desired exposure level, using a combination of different apertures and focal plane locations. The individual photographs are composited together using a variant of the Photomontage method [18] to create a large DOF composite. As a by-product, they also get a coarse depth map of the scene. Our approach does not recover scene depth, but can produce an all-in-focus photograph from a single, well-exposed image.

There is similar work on moving the detector during



(a) (b) (c)

Fig. 1. (a) A scene point M, at a distance u from the lens, is imaged in perfect focus by a detector at a distance v from the lens. If the detector is shifted to a distance p from the lens, M is imaged as a blurred circle with diameter b centered around m'. (b) Our flexible DOF camera translates the detector along the optical axis during the integration time of an image. By controlling the starting position, speed, and acceleration of the detector, we can manipulate the DOF in powerful ways. (c) Our prototype flexible DOF camera.

image integration [19]. However, their focus is on handling motion blur, for which they propose to move the detector *perpendicular* to the optical axis. Some previous works have also varied the orientation or location of the image detector. Krishnan and Ahuja [3] tilt the detector and capture a panoramic image sequence, from which they compute an all-focused panorama and a depth map. For video super-resolution, Ben-Ezra et al. [20] capture a video sequence by instantaneously shifting the detector within the image plane, in between the integration periods of successive video frames.

Recently, it has been shown that a detector with a rolling shutter can be used to estimate the pose and velocity of a fast moving object [21]. We show how a rolling shutter detector can be used to focus on tilted scene planes as well as non-planar scene surfaces.

3 CAMERA WITH PROGRAMMABLE DEPTH OF FIELD

Consider Figure 1(a), where the detector is at a distance v from a lens with focal length f and an aperture of diameter a. A scene point M is imaged in perfect focus at m, if its distance u from the lens satisfies the lens law:

$$\frac{1}{f} = \frac{1}{u} + \frac{1}{v}. \tag{1}$$

As shown in the figure, if the detector is shifted to a distance p from the lens (dotted line), M is imaged as a blurred circle (the circle of confusion) centered around m'. The diameter b of this circle is given by

$$b = \frac{a}{v}|(v-p)|\,. \tag{2}$$

The distribution of light energy within the blur circle is referred to as the point spread function (PSF). The PSF can be denoted as $P(r,u,p)$, where r is the distance of an image point from the center m' of the blur circle. An idealized model for characterizing the PSF is the pillbox function:

$$P(r,u,p) = \frac{4}{\pi b^2}\Pi(\frac{r}{b}), \tag{3}$$

where, $\Pi(x)$ is the rectangle function, which has a value 1, if $|x| < 1/2$ and 0 otherwise. In the presence of optical aberrations, the PSF deviates from the pillbox function and is then often modeled as a Gaussian function:

$$P(r,u,p) = \frac{2}{\pi(gb)^2}\exp(-\frac{2r^2}{(gb)^2}), \tag{4}$$

where g is a constant.

We now analyze the effect of moving the detector during an image's integration time. For simplicity, consider the case where the detector is translated along the optical axis, as in Figure 1(b). Let $p(t)$ denote the detector's distance from the lens as a function of time. Then the aggregate PSF for a scene point at a distance u from the lens, referred to as the *integrated PSF* (IPSF), is given by

$$IP(r,u) = \int_0^T P(r,u,p(t))\ \ dt, \tag{5}$$

where T is the total integration time. By programming the detector motion $p(t)$ – its starting position, speed, and acceleration – we can change the properties of the resulting IPSF. This corresponds to sweeping the focal plane through the scene in different ways. The above analysis only considers the translation of the detector along the optical axis (as implemented in our prototype camera). However, this analysis can be easily extended to more general detector motions, where both its position and orientation are varied during image integration.

Figure 1(c) shows our flexible DOF camera. It consists of a 1/3" Sony CCD (with 1024x768 pixels) mounted on a Physik Instrumente M-111.1DG translation stage. This stage has a DC motor actuator that can translate the detector through a 15 mm range at a top speed of 2.7 mm/sec and can position it with an accuracy of 0.05 microns. The translation direction is along the optical axis of the lens. The CCD shown has a global shutter and was used to implement extended DOF and discontinuous DOF. For realizing tilted and non-planar DOFs, we used a 1/2.5" Micron CMOS detector (with 2592x1944 pixels) which has a rolling shutter.

Lens Focal Length	Scene Depth Range	Required Detector Translation	Maximum Change in Image Position
9.0mm	1m - ∞	81.7 μm	4.5 pixels
	.5m - ∞	164.9 μm	5.0 pixels
	.2m - 0.5m	259.1 μm	7.2 pixels
12.5mm	1m - ∞	158.2 μm	3.6 pixels
	.5m - ∞	320.5 μm	5.6 pixels
	.2m - 0.5m	512.8 μm	8.5 pixels

Fig. 2. Translation of the detector required for sweeping the focal plane through different scene depth ranges. The maximum change in the image position of a scene point that results from this translation, when a 1024x768 pixel detector is used, is also shown.

The table in Figure 2 shows detector translations (third column) required to sweep the focal plane through various depth ranges (second column), using lenses with two different focal lengths (first column). As we can see, the detector has to be moved by very small distances to sweep very large depth ranges. Using commercially available micro-actuators, such translations are easily achieved within typical image integration times (a few milliseconds to a few seconds).

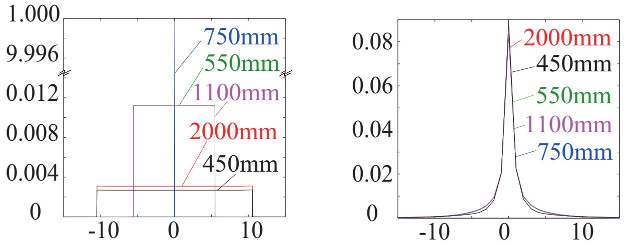
It must be noted that when the detector is translated, the magnification of the imaging system changes [1]. The fourth column of the table in Figure 2 lists the maximum change in the image position of a scene point for different translations of a 1024x768 pixel detector. For the detector motions we require, these changes in magnification are very small. This does result in the images not being perspectively correct, but the distortions are imperceptible. More importantly, the IPSFs are not significantly affected by such a magnification change, since a scene point will be in high focus only for a small fraction of this change and will be highly blurred over the rest of it. We verify this in the next section.

4 EXTENDED DEPTH OF FIELD (EDOF)

In this section, we show that we can capture scenes with EDOF by translating a detector with a global shutter at a constant speed during image integration. We first show that the IPSF for an EDOF camera is nearly invariant to scene depth for all depths swept by the focal plane. As a result, we can deconvolve a captured image with the IPSF to obtain an image with EDOF and high SNR.

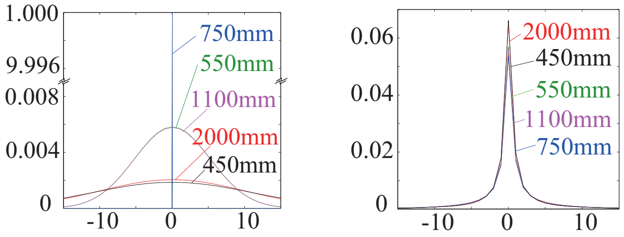
4.1 Depth Invariance of IPSF

Consider a detector translating along the optical axis with constant speed s, i.e., $p(t) = p(0) + st$. If we assume



(a) Normal Camera PSF (Pillbox)

(b) EDOF Camera IPSF



(c) Normal Camera PSF (Gaussian)

(d) EDOF Camera IPSF

Fig. 3. Simulated (a,c) normal camera PSFs and (b,d) EDOF camera IPSFs, obtained using pillbox and Gaussian lens PSF models for 5 scene depths. Note that the IPSFs are almost invariant to scene depth.

that the PSF of the lens can be modeled using the pillbox function in Equation 3, the IPSF in Equation 5 becomes

$$IP(r,u) = \frac{uf}{(u-f)\pi asT}\left(\frac{\lambda_0+\lambda_T}{r} - \frac{2\lambda_0}{b(0)} - \frac{2\lambda_T}{b(T)}\right), \quad (6)$$

where, $b(t)$ is the blur circle diameter at time t, and $\lambda_t = 1$ if $b(t) \geq 2r$ and 0 otherwise. If we use the Gaussian function in Equation 4 for the lens PSF, we get

$$IP(r,u) = \frac{uf}{(u-f)\sqrt{2\pi}rasT}\left(\operatorname{erfc}\left(\frac{r}{\sqrt{2}gb(0)}\right) + \operatorname{erfc}\left(\frac{r}{\sqrt{2}gb(T)}\right)\right). \quad (7)$$

Figures 3(a) and (c) show 1D profiles of a normal camera's PSFs for 5 scene points with depths between 450 and 2000 mm from a lens with focal length $f = 12.5$ mm and $f/\# = 1.4$, computed using Equations 3 and 4 (with $g = 1$), respectively. In this simulation, the normal camera was focused at a distance of 750 mm. Figures 3(b) and (d) show the corresponding IPSFs of an EDOF camera with the same lens, $p(0) = 12.5$ mm, $s = 1$ mm/sec, and $T = 0.36$ sec, computed using Equations 6 and 7, respectively. As expected, the normal camera's PSF varies dramatically with scene depth. In contrast, the IPSFs of the EDOF camera derived using both pillbox and Gaussian PSF models look almost identical for all 5 scene depths, i.e., *the IPSFs are depth invariant*.

The above analysis is valid for a scene point that projects to the center pixel in the image. For any other scene point, due to varying magnification, the location of its image will change during image integration (see

---

[1] Magnification is defined as the ratio of the distance between the lens and the detector and the distance between the lens and the object. By translating the detector we are changing the distance between the lens and the detector, and hence changing the magnification of the system during image integration.

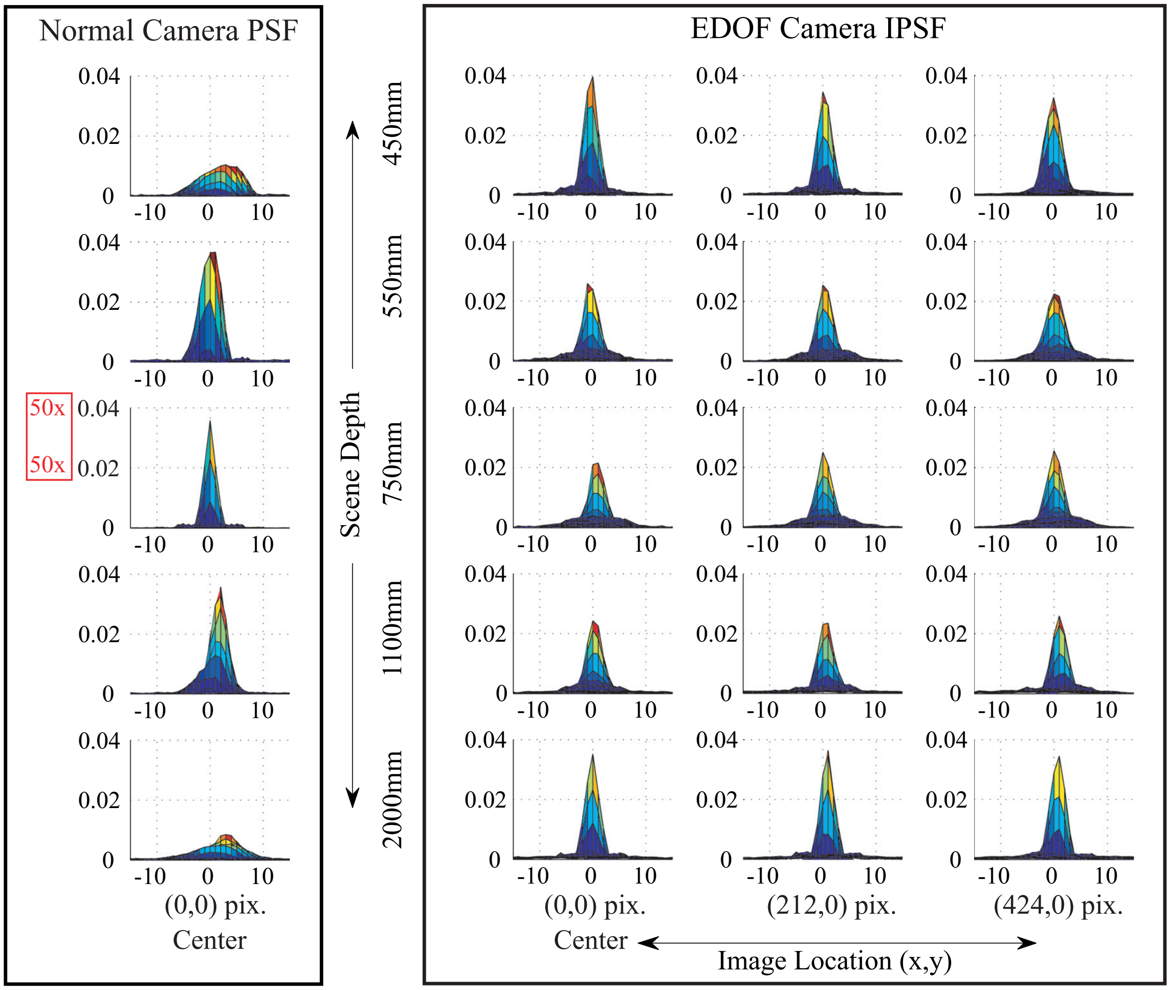


Fig. 4. (Left column) The measured PSF of a normal camera shown for 5 different scene depths. Note that the scale of the plot in the center row is 50 times that of the other plots. (Right columns) The measured IPSF of our EDOF camera shown for different scene depths (vertical axis) and image locations (horizontal axis). The EDOF camera's IPSFs are almost invariant to scene depth and image location.

Figure 2). However, a scene point will be in high focus for only a short duration of this change (contributing to the peak of the PSF), and be highly blurred the rest of the time (contributing to the tail of the PSF). As a result, the approximate depth invariance of the PSF can be expected to hold over the entire image.

To empirically verify this, we measured a normal camera's PSFs and the EDOF camera's IPSFs for several scene depths, by capturing images of small dots placed at different depths. Both cameras have $f = 12.5$ mm, $f/\# = 1.4$, and $T = 0.36$ sec. The detector motion parameters for the EDOF camera are $p(0) = 12.5$ mm and $s = 1$ mm/sec. This corresponds to sweeping the focal plane from ∞ to 446.53 mm. The first column of Figure 4 shows the measured PSF at the center pixel of the normal camera for 5 scene depths; the camera was focused at a distance of 750 mm. (Note that the scale of the plot in the center row is 50 times that of the other plots.) Columns 2-4 of the figure show the IPSFs of the EDOF camera for 5 different scene depths and 3 different image locations. We can see that, while the normal camera's PSFs vary widely with scene depth, the EDOF camera's IPSFs appear almost invariant to both scene depth and spatial location. This also validates our claim that the small magnification changes that arise due to detector motion do not have a significant impact on the IPSFs.

To quantitatively analyze the depth and space invariances of the IPSF, we use the Wiener reconstruction error when an image is blurred with kernel k_1 and then deconvolved with kernel k_2. In order to account for the fact that in natural images all frequencies do not have the same importance, we weigh this reconstruction error to get the following measure of dissimilarity of two PSFs k_1 and k_2

$$d(k_1, k_2) = \sum_{\omega}\left(\frac{|K_1(\omega) - K_2(\omega)|^2}{|K_1(\omega)|^2 + \epsilon} + \frac{|K_1(\omega) - K_2(\omega)|^2}{|K_2(\omega)|^2 + \epsilon}\right) \ |F(\omega)|^2, \tag{8}$$

where, K_i is the Fourier transform of k_i, ω represents

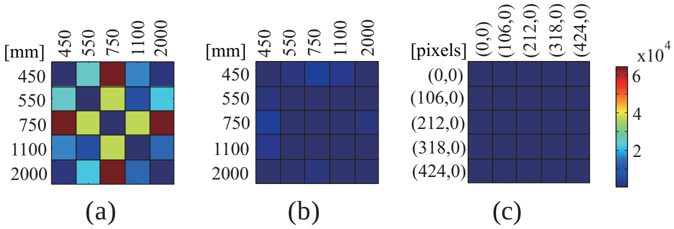


Fig. 5. (a) Pair-wise dissimilarity of a normal camera's measured PSFs at the center pixel for 5 scene depths. The camera was focused at a distance of 750 mm. (b) Pair-wise dissimilarity of the EDOF camera's measured IPSFs at the center pixel for 5 scene depths. (c) Pair-wise dissimilarity of the EDOF camera's measured IPSFs at 5 different locations along the center row of the image, for scene points at a distance of 750 mm. (0,0) denotes the center of the image.

2D frequency, $|F|^2$ is a weighting term that encodes the power fall-off of Fourier coefficients in natural images [22], and ϵ is a small positive constant that ensures that the denominator terms are non-zero. Figure 5(a) shows a visualization of the pair-wise dissimilarity between the normal camera's PSFs measured at the center pixel, for 5 different scene depths. Figure 5(b) shows a similar plot for the EDOF camera's IPSFs measured at the center pixel, while Figure 5(c) shows the pair-wise dissimilarity of the IPSFs at 5 different image locations but for the same scene depth. These plots further illustrate the invariance of an EDOF camera's IPSF. We have empirically observed that the approximate invariance holds reasonably well for the entire range of scene depths swept by the focal plane during the detector's motion. However, the invariance is slightly worse for depths that correspond to roughly 10% of the distance traveled by the detector at both the beginning and end of its motion.

Hausler's work [1] describes how depth independent blur can be realized for object-side telecentric imaging systems. The above results demonstrate that changing the focus by moving the detector at a constant velocity, during image integration, yields approximate depth independent blur for conventional imaging geometries. However, it should be noted that Hausler's analysis is more rigorous – he uses a more accurate model for defocus [23] than ours. Also, by virtue of the imaging system being object-side telecentric, his analysis did not have to model magnification. In our approach, though magnification changes during image integration, we have not explicitly modeled its effects.

4.2 Computing EDOF Images using Deconvolution

Since the EDOF camera's IPSF is invariant to scene depth and image location, we can deconvolve a captured image with a single IPSF to get an image with greater DOF. A number of techniques have been proposed for deconvolution, Richardson-Lucy and Wiener [24] being two popular ones. For our results, we have used the approach of Dabov et al. [25], which combines Wiener deconvolution and block-based denoising. In all our experiments, we used the IPSF shown in the first row and second column of Figure 4 for deconvolution.

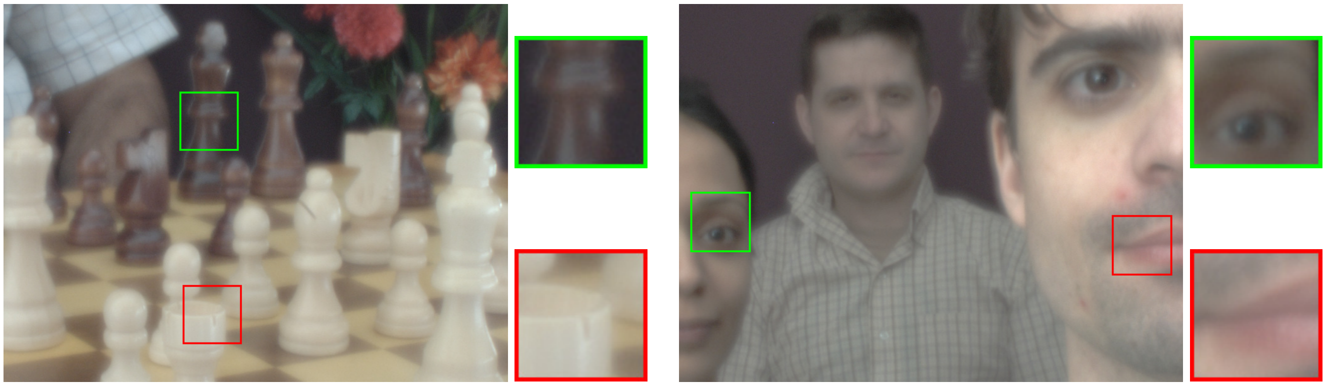
Figures 6(a) show images captured by our EDOF camera. They were captured with a 12.5 mm Fujinon lens with $f/1.4$ and 0.36 second exposures. Notice that the captured images look slightly blurry, but high frequencies of all scene elements are captured. These scenes span a depth range of approximately 450 mm to 2000 mm – 10 times larger than the DOF of a normal camera with identical lens settings. Figures 6(b) show the EDOF images computed from the captured images, in which all scene elements appear focused[2]. Figures 6(c) show images captured by a normal camera with the same $f/\#$ and exposure time. Scene elements at the center depth are in focus. We can get a large DOF image using a smaller aperture. Images captured by a normal camera with the same exposure time, but with a smaller aperture of $f/8$ are shown in Figures 6(d). The intensities of these images were scaled up so that their dynamic range matches that of the corresponding computed EDOF images. All scene elements look reasonably sharp, but the images are very noisy as can be seen in the insets (zoomed). The computed EDOF images have much less noise, while having comparable sharpness, i.e. our EDOF camera can capture scenes with large DOFs as well as high SNR. Figure 7 shows another example, of a scene captured outdoors at night. As we can see, in a normal camera, the tradeoff between DOF and SNR is extreme for such dimly lit scenes. Our EDOF camera operating with a large aperture is able to capture something in this scene, while a normal camera with a comparable DOF is too noisy to be useful. Several denoising algorithms have been proposed and its conceivable that they can be used to improve the appearance of images captured with a small aperture, like the images in Figure 6(d). However, it is unlikely that they can be used to restore images like the one in Figure 7(d). High resolution versions of these images as well as other examples can be seen at [27].

Since we translate the detector at a constant speed, the IPSF does not depend on the direction of motion – it is the same whether the detector moves from a distance a from the lens to a distance b from the lens or from a distance b from the lens to a distance a from the lens. We can exploit this to get EDOF video by moving the detector alternately forward one frame and backward the next. Figure 8(a) shows a frame from a video sequence captured in this fashion and Figure 8(b) shows the EDOF frame computed from it. In this example, we were restricted by the capabilities of our actuator and were able to achieve only 1.5 frames/sec. To reduce motion blur, the camera was placed on a slowly moving robotic arm. For comparison, Figures 8(c) and (d) show frames from video sequences captured by a

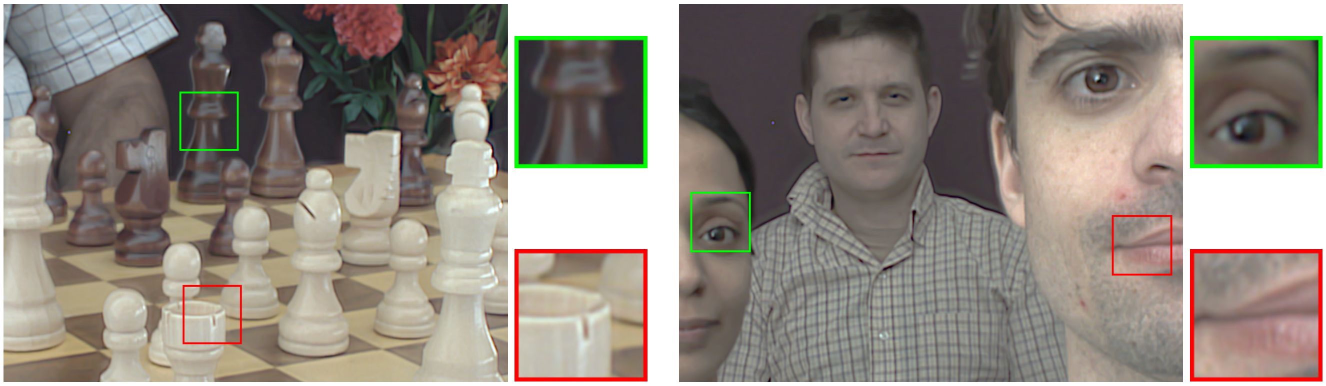
---

[2] The computed EDOF images do have artifacts, like ringing, that are typical of deconvolution [26].

Please zoom in to see noise and defocus blur

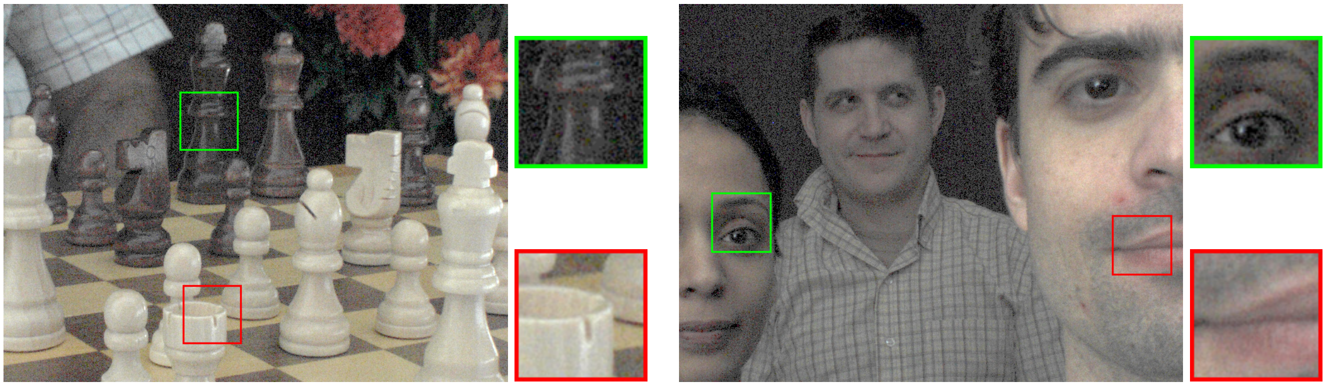


(a) Images Captured by our EDOF Camera ($f/1.4$)



(b) EDOF Images computed from the Captured Images

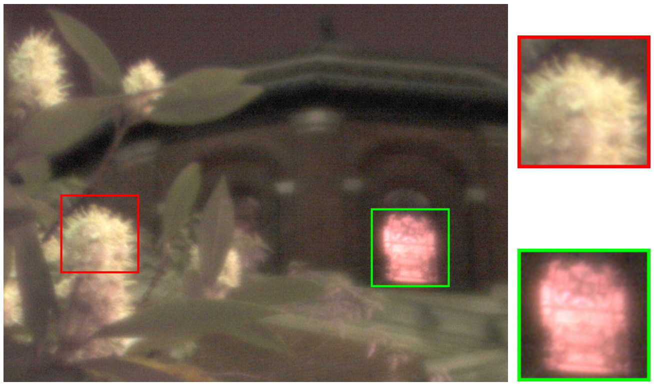
(c) Images Captured by a Normal Camera ($f/1.4$, Center Focus)



(d) Images Captured by a Normal Camera ($f/8$, Center Focus) with Scaling

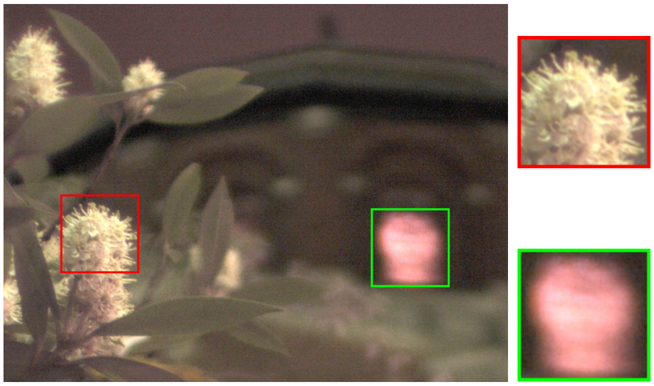
Fig. 6. Images were captured with an exposure time of 0.36 seconds.

Please zoom in to see noise and defocus blur

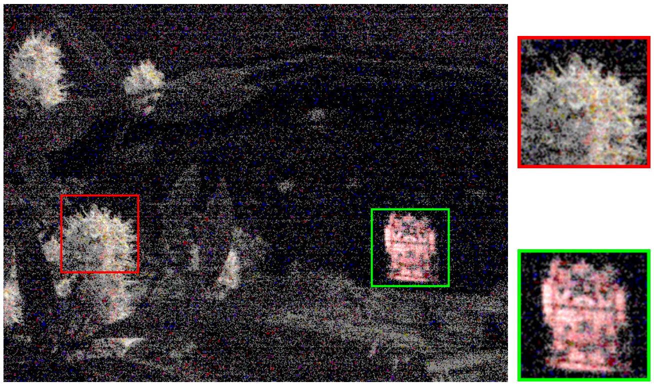


(a) Image captured by our EDOF camera ($f/1.4$)

(b) Computed EDOF Image



(c) Image from Normal Camera ($f/1.4$, Near Focus)



(d) Image from Normal Camera ($f/8$, Near Focus) with Scaling

Fig. 7. Images were captured with an exposure time of 0.72 seconds.

normal camera operating at $f/1.4$ and $f/8$ respectively.

4.3 Analysis of SNR Benefits of EDOF Camera

We now analyze the SNR benefits of using our approach to capture scenes with extended DOF. Deconvolution using Dabov et al.'s method [25] produces visually appealing results, but since it has a non-linear denoising step, it is not suitable for analyzing the SNR of deconvolved captured images. Therefore, we performed a simulation that uses Wiener deconvolution [24]. Given an IPSF k, we convolve it with a natural image I, and add zero-mean white Gaussian noise with standard deviation σ. The resulting image is then deconvolved with k to get the EDOF image $\hat{I}$. The standard deviation $\hat{\sigma}$ of $(I - \hat{I})$ is a measure of the noise in the deconvolution result when the captured image has noise σ.

The degree to which deconvolution amplifies noise depends on how much the high frequencies are attenuated by the IPSF. This, in turn, depends on the distance through which the detector moves during image integration – as the distance increases, so does the attenuation of high frequencies. This is illustrated in Figure 9(a), which shows (in red) the MTF (magnitude of the Fourier transform) for a simulated IPSF k_1, derived using the pillbox lens PSF model. In this case, we use the same detector translation and parameters as in our EDOF experiments (Section 4.1). The MTF of the IPSF k_2 obtained when the detector translation is halved (keeping the mid-point of the translation the same) is also shown (in blue). As expected, k_2 attenuates the high frequencies less than k_1.

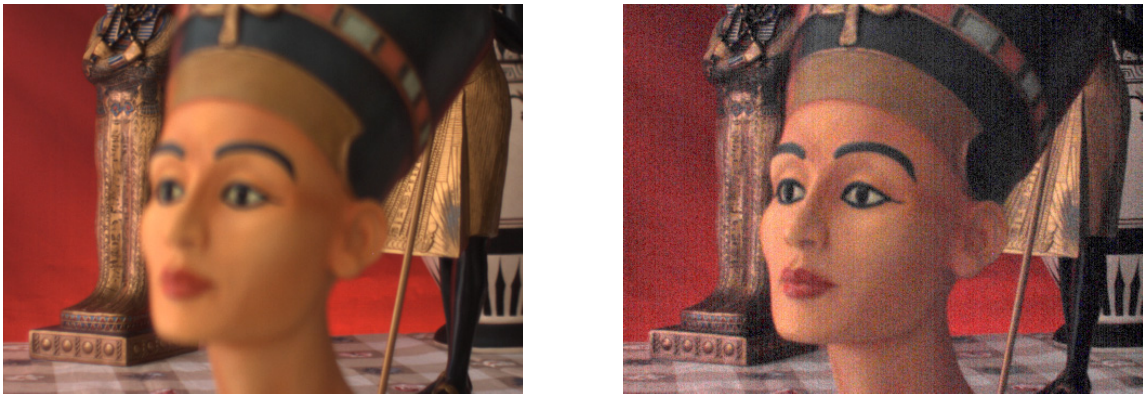
We analyzed the SNR benefits for these two IPSFs for different noise levels in the captured image. The table in Figure 9(b) shows the noise produced by a normal camera for different aperture sizes, given the noise level for the largest aperture, $f/1.4$. (Image brightness is assumed to lie between 0 and 1.) The last two rows show the effective noise levels for EDOF cameras with IPSFs k_1 and k_2, respectively. The last column shows the effective DOFs realized; the normal camera was assumed to be focused at a distance of 750 mm with a maximum permissible circle of confusion of 14.1 μm for a 1/3" sensor. Note that, as the noise level in the captured image increases, the SNR benefits of EDOF cameras increase[3]. As an example, if the noise of a normal camera at $f/1.4$ is 0.01, then the EDOF camera with IPSF k_1 has the SNR of a normal camera operating at $f/2.8$, but has a DOF that is greater than that of a normal camera at $f/8$.

In the above analysis, the SNR was averaged over all frequencies. However, it must be noted that SNR is frequency dependent - SNR is greater for lower frequencies than for higher frequencies in the deconvolved EDOF

---

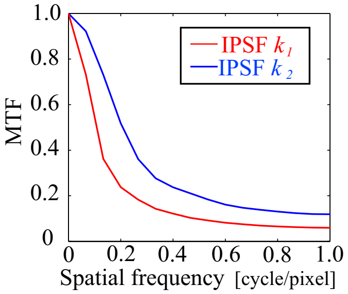
[3] For low noise levels, instead of capturing a well exposed image with an EDOF camera, one could possibly use a normal camera to capture multiple images with very low exposures (so that the total exposure time is the same) and form a focal stack like [14], [15], [16], [17], provided the camera is able to change focus and capture the images fast enough. However, as the noise level in captured images increases, the SNR benefits of EDOF cameras clearly increase.

(a) Video frame captured by our EDOF camera ($f/1.4$) (b) Computed EDOF Frame



(c) Video frame from Normal Camera ($f/1.4$) (d) Video frame from Normal Camera ($f/8$) with Scaling

Fig. 8. An example that demonstrates how our approach can be used to capture EDOF video and its benefits over a normal camera. These videos can be seen at [27].



(a)

Camera	f/ #	Noise standard deviation					DOF (mm)
Normal	1.4	0.001	0.002	0.005	0.010	0.020	140.98
Normal	2.8	0.004	0.008	0.020	0.040	0.080	289.57
Normal	4	0.008	0.016	0.041	0.082	0.163	429.77
Normal	5.6	0.016	0.032	0.080	0.160	0.320	649.21
Normal	8	0.033	0.065	0.163	0.327	0.653	1114.56
EDOF(k_1)	1.4	0.013	0.021	0.036	0.049	0.062	1622.44
EDOF(k_2)	1.4	0.0073	0.0133	0.0262	0.0399	0.0570	528.56

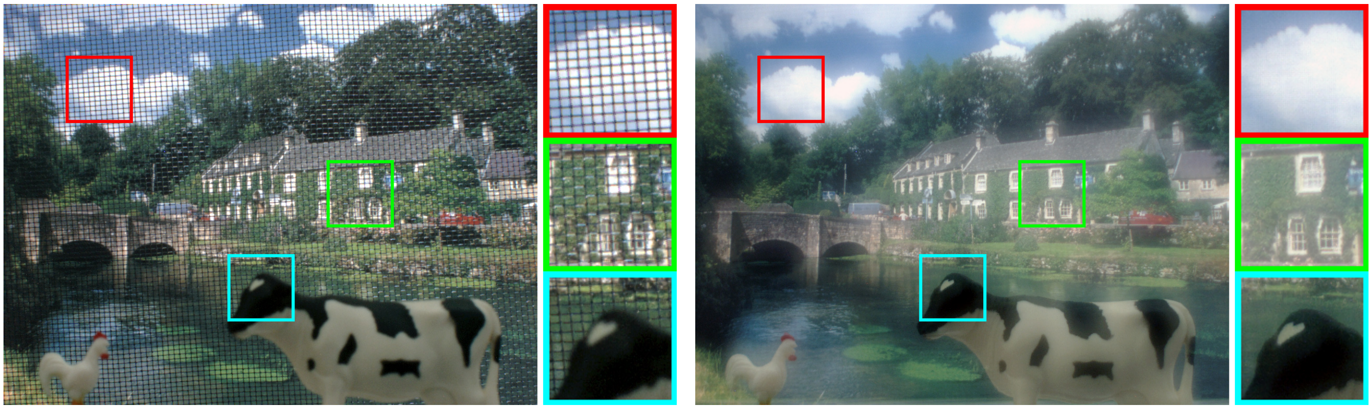
(b)

Fig. 9. (a) MTFs of simulated IPSFs, k_1 and k_2, of an EDOF camera corresponding to the detector traveling two different distances during image integration. (b) Comparison of effective noise and DOF of a normal camera and a EDOF camera with IPSFs k_1 and k_2. The image noise of a normal camera operating at $f/1.4$ is assumed to be known.

images. Hence, high frequencies in an EDOF image would be degraded, compared to the high frequencies in a perfectly focused image. However, in our experiments this degradation is not strong, as can be seen in the insets of Figure 6 and the full resolution images at [27]. Different frequencies in the image having different SNRs illustrates the tradeoff that our EDOF camera makes. In the presence of noise, instead of capturing with high fidelity, high frequencies over a small range of scene depths (the depth of field of a normal camera), our EDOF camera captures with slightly lower fidelity, high frequencies over a large range of scene depths.

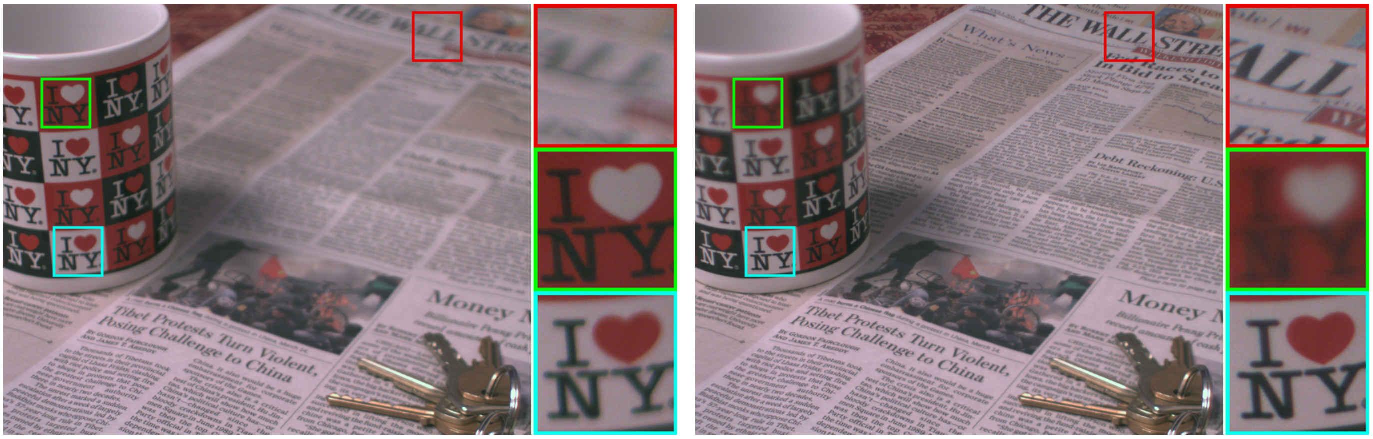
5 DISCONTINUOUS DEPTH OF FIELD

Consider the image in Figure 10(a), which shows two toys (cow and hen) in front of a scenic backdrop with a wire mesh in between. A normal camera with a small DOF can capture either the toys or the backdrop in focus,



(a) Image from Normal Camera (b) Image from Our Camera ($f/1.4$)

Fig. 10. (a) An image captured by a normal camera with a large DOF. (b) An image captured by our flexible DOF camera, where the toy cow and hen in the foreground and the landscape in the background appear focused, while the wire mesh in between is optically erased via defocusing.



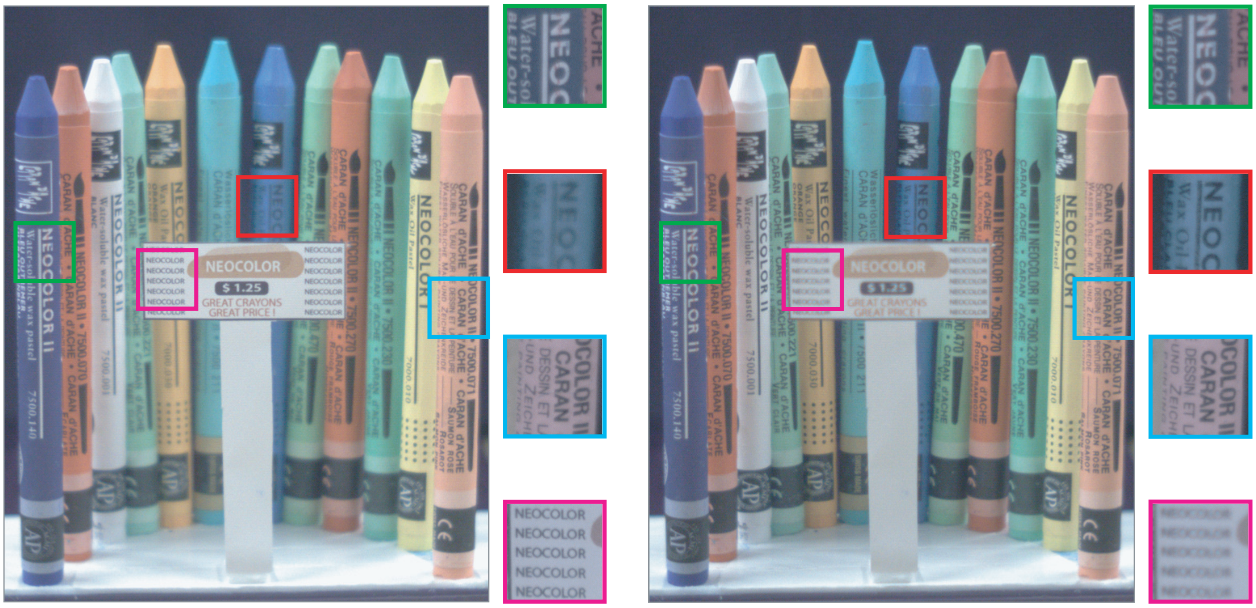
(a) Image from Normal Camera ($f/1.4$, T=0.03sec) (b) Image from our Camera ($f/1.4$, T=0.03sec)

Fig. 11. (a) An image captured by a normal camera of a table top inclined at 53° with respect to the lens plane. (b) An image captured by our flexible DOF camera, where the DOF is tilted by 53°. The entire table top (with the newspaper and keys) appears focused. Observe that the top of the mug is defocused, but the bottom appears focused, illustrating that the focal plane is aligned with the table top. Three scene regions of both the images are shown at a higher resolution to highlight the defocus effects.

while eliminating the mesh via defocusing. However, since its DOF is a single continuous volume, it cannot capture both the toys and the backdrop in focus and at the same time eliminate the mesh. If we use a large aperture and program our camera's detector motion such that it first focuses on the toys for a part of the integration time, and then moves quickly to another location to focus on the backdrop for the remaining integration time, we obtain the image in Figure 10(b). While this image includes some blurring, it captures the high frequencies in two disconnected DOFs - the foreground and the background - but almost completely eliminates the wire mesh in between. This is achieved without any post-processing. Note that we are not limited to two disconnected DOFs; by pausing the detector at several locations during image integration, more complex DOFs can be realized.

6 TILTED DEPTH OF FIELD

Normal cameras can focus on only fronto-parallel scene planes. On the other hand, view cameras [2], [3] can be made to focus on tilted scene planes by adjusting the orientation of the lens with respect to the detector. We show that our flexible DOF camera can be programmed to focus on tilted scene planes by simply translating (as in previous applications) a detector with a rolling electronic shutter. A large fraction of CMOS detectors are of this type – while all pixels have the same integration time, successive rows of pixels are exposed with a slight time lag. If the exposure time is sufficiently small, then upto an approximation, we can say that the different rows of the image are exposed independently. When such a detector is translated with uniform speed s, during the frame read out time T of an image, we emulate a tilted image detector. If this tilted detector makes an angle θ



(a) Image from Normal Camera ($f/1.4, T = 0.01sec$) (b) Image from our Camera ($f/1.4, T = 0.01sec$)

Fig. 12. (a) An image captured by a normal camera of crayons arranged on a semi-circle with a price tag in the middle placed at the same depth as the left-most and right-most crayons. Only the price tag and the extreme crayons are in focus. (b) An image captured by our flexible DOF camera where the DOF is curved to be aligned with the crayons – all the crayons are in focus, while the price tag is defocused. Four scene regions of both the images are shown at a higher resolution to highlight the defocus effects.

with the lens plane, then the focal plane in the scene makes an angle ϕ with the lens plane, where θ and ϕ are related by the well-known Scheimpflug condition [4]:

$$\begin{aligned} \theta &= \tan^{-1}\left(\frac{sT}{H}\right) \text{ and} \\ \phi &= \tan^{-1}\left(\frac{2f\tan(\theta)}{2p(0) + H\tan(\theta) - 2f}\right). \end{aligned} \qquad (9)$$

Here, H is the height of the detector. Therefore, by controlling the speed s of the detector, we can vary the tilt angle of the image detector, and hence the tilt of the focal plane and its associated DOF.

Figure 11 shows a scene where the dominant scene plane – a table top with a newspaper, keys and a mug on it – is inclined at an angle of approximately 53° with the lens plane. As a result, a normal camera is unable to focus on the entire plane, as seen in Figure 11(a). By translating a rolling-shutter detector (1/2.5” CMOS sensor with a 70msec exposure lag between the first and last row of pixels) at 2.7 mm/sec, we emulate a detector tilt of 2.6°. This enables us to achieve the desired DOF tilt of 53° (from Equation 9) and capture the table top (with the newspaper and keys) in focus, as shown in Figure 11(b). Observe that the top of the mug is not in focus, but the bottom appears focused, illustrating the fact that the DOF is tilted to be aligned with the table top. Note that there is no post-processing here.

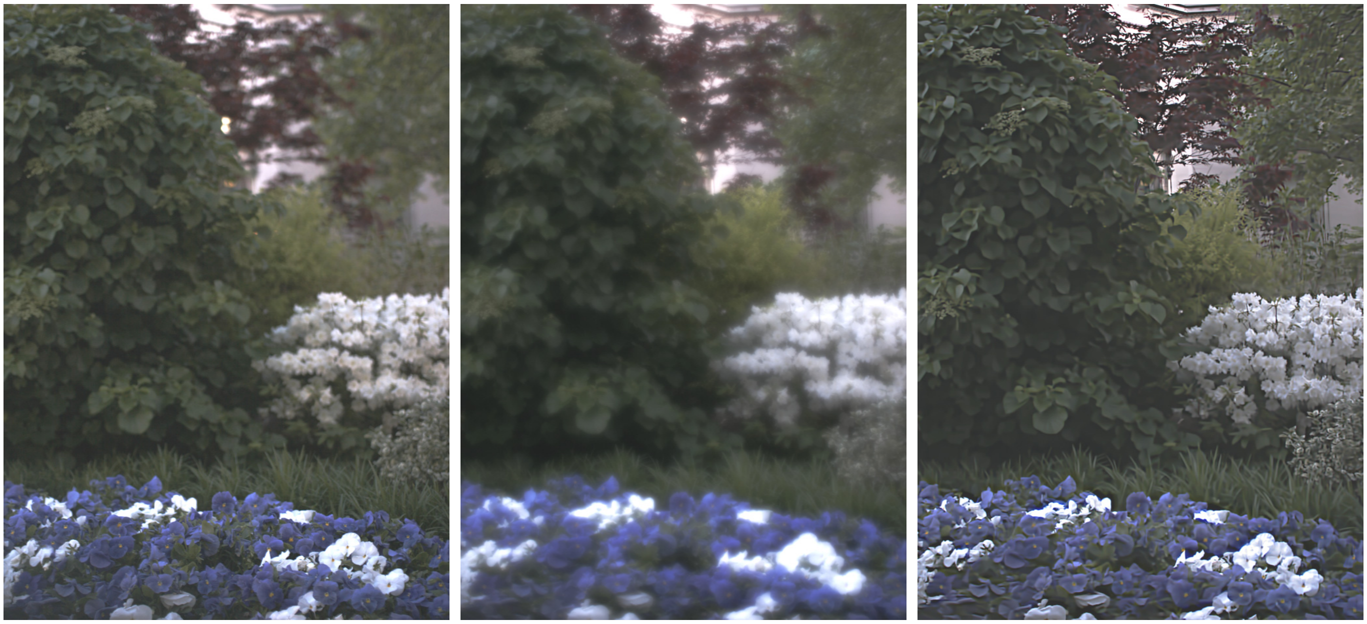
7 NON-PLANAR DEPTH OF FIELD

In the previous section, we have seen that by uniformly translating a detector with a rolling shutter we can emulate a tilted image detector. Taking this idea forward, if we translate such a detector in some non-uniform fashion (varying speed), we can emulate a non-planar image detector. Consequently, we get a non-planar focal surface and hence a non-planar DOF. This is in contrast to a normal camera which has a planar focal surface and whose DOF is a fronto-parallel slab.

Figure 12 (a) shows a scene captured by a normal camera. It has crayons arranged on a semi-circle with a price tag in the middle placed at the same depth as the left-most and right-most crayons. Only the two extreme crayons on either side and the price tag are in focus; the remaining crayons are defocused. Say, we want to capture this scene so that the DOF is ‘curved’ – the crayons are in focus while the price tag is defocused. We set up a non-uniform motion of the detector to achieve this desired DOF, which can be seen in Figure 12 (b).

8 EXPLOIT CAMERA’S FOCUSING MECHANISM TO MANIPULATE DEPTH OF FIELD

Till now we have seen that by moving the detector during image integration, we can manipulate the DOF. However, it must be noted that whatever effect we get by moving the detector, we can get exactly the same effect by moving the lens (in the opposite direction). In fact, cameras already have mechanisms to do this; this is what happens during focusing. Hence, we can exploit the camera’s focusing mechanism to manipulate



(a) Image from Normal Camera ($f/1.4, T = 0.6sec$)

(b) Image captured by rotating the focus ring ($f/1.4, T = 0.6sec$)

(c) Computed EDOF Image

Fig. 13. (a) Image captured by a Canon EOS 20D SLR camera with a Sigma 30 mm lens operating at $f/1.4$, where only the near flowers are in focus. (b) Image captured by the camera when the focus ring was manually rotated uniformly during image integration. (c) Image with extended DOF computed from the image in (b).

DOF. Figure 13(a) shows an image captured by a normal SLR camera (Canon EOS 20D with a Sigma 30 mm lens) at $f/1.4$, where only the near flowers are in focus. To capture this scene with an extended DOF, we manually rotated the focus ring of the SLR camera lens uniformly during image integration. For the lens we used, uniform rotation corresponds to moving the lens at a roughly constant speed. Figure 13(b) shows an image captured in this fashion. Figure 13(c) shows the EDOF image computed from it, in which the entire scene appears sharp and well focused. For deconvolution, we used the analytic PSF given by Equation 6. These images as well as other examples can be seen at [27].

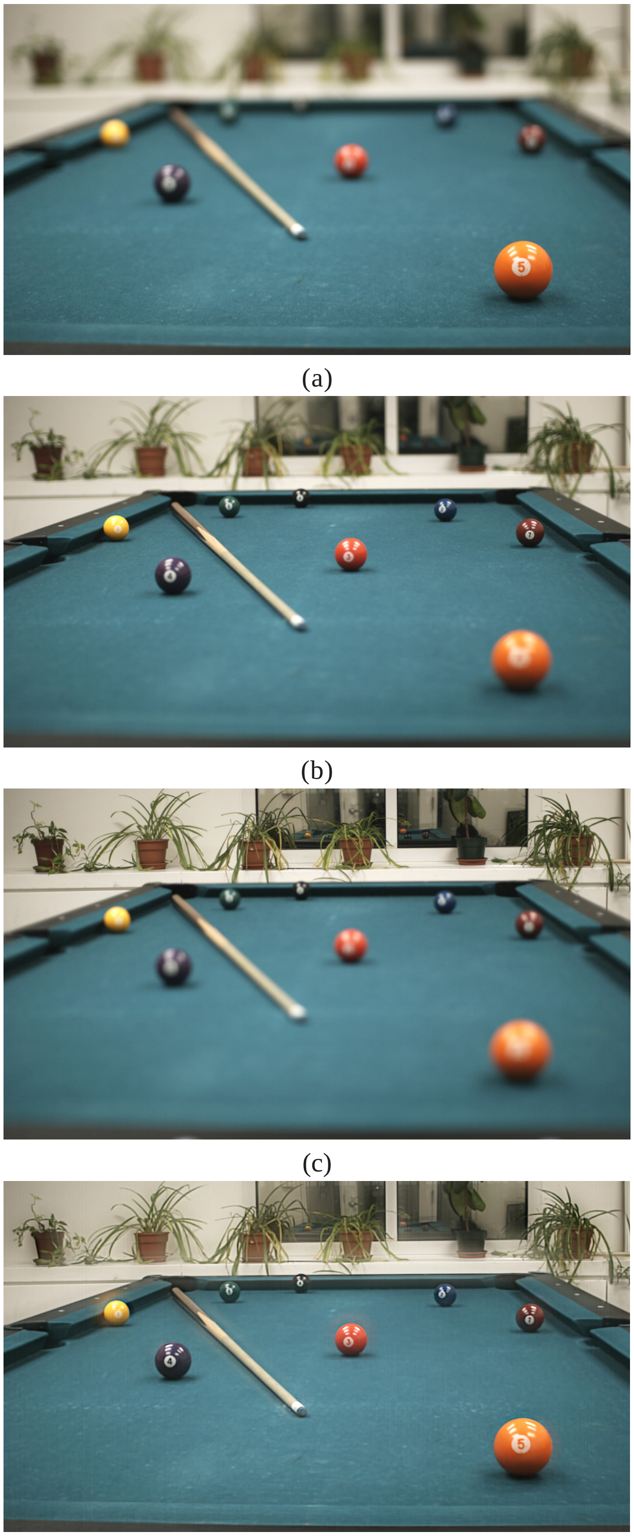
9 COMPUTING AN ALL-FOCUSED IMAGE FROM A FOCAL STACK

Our approach to extended DOF also provides a convenient means to compute an all-focused image from a focal stack. Traditionally, given a focal stack, for every pixel we have to determine in which image that particular pixel is in-focus [28], [29]. Some previous works have tackled this as a labeling problem, where the label for every pixel is the image where the pixel is in-focus. The labels are optimized using a Markov Random Field that is biased towards piece-wise smoothness [18], [17].

We propose an alternate approach that leverages our observations in Section 4.1. We propose to compute a weighted average of all the images of the focal stack (compensating for magnification effects if possible), where the weights are chosen to mimic changing the distance between the lens and the detector at a constant speed. From Section 4.1 we know that this average image would have approximately depth independent blur. Hence, deconvolution with a single blur kernel will give a sharp image in which all scene elements appear focused. Figures 14(a,b,c) show three of the 28 images that form a focal stack. These were captured with a Canon 20D SLR camera with a Sigma 30 mm lens operating at $f/1.4$. Figure 14(d) shows the all-focused image computed from the focal stack using this approach. We are not claiming that this technique is the best for computing an all focused image from a focal stack. As noted earlier, deconvolution artifacts could appear in the resulting images and high frequencies would be captured with lower fidelity. This example illustrates how our approach can be used to realize a simpler (possibly slightly inferior) solution to this problem than conventional approaches.

10 DISCUSSION

In this paper we have proposed a camera with a flexible DOF. DOF is manipulated in various ways by controlling the motion of the detector during image integration. We have shown how such a system can capture scenes with extended DOF while using large apertures. We have also shown that we can create DOFs that span multiple disconnected volumes. In addition, we have demonstrated that our camera can focus on tilted scene planes as well as non-planar scene surfaces. Finally, we have shown that we can manipulate DOF by exploiting the focusing



(a)

(b)

(c)

(d)

Fig. 14. (a,b,c) Three out of 28 images that form a focal stack. The images were captured with a Canon 20D camera with a Sigma 30 mm lens operating at $f/1.4$. (d) The all-focused image computed from the focal stack images using the approach described in Section 9.

mechanism of the lens. This can be very convenient and practical, especially for camera manufacturers.

Effects at Occlusion Boundaries: For our EDOF camera, we have not explicitly modeled the defocus effects at occlusion boundaries. Due to defocus blur, image points that lie close to occlusion boundaries can receive light from scene points at very different depths. However, since the IPSF of the EDOF camera is nearly depth invariant, the aggregate IPSF for such an image point can be expected to be similar to the IPSF of points far from occlusion boundaries. In some of our experiments, we have seen artifacts at occlusion boundaries. These can possibly be eliminated using more sophisticated deconvolution algorithms such as [26], [30]. In the future, we would like to analyze in detail the effects at occlusion boundaries, similar to works like [31] and [32]. Note that in tilted and non-planar DOF examples occlusion boundaries are correctly captured; there are no artifacts.

Effects of Scene Motion: The simple off-the-shelf actuator that we used in our prototype has low translation speeds and so we had to use exposure times of about $1/3^{rd}$ of a second to capture EDOF images. However, we have not observed any visible artifacts in EDOF images computed for scenes with typical object motion (see Figure 6). With faster actuators, like piezoelectric stacks, exposure times can be made much smaller and thereby allow captured scenes to be more dynamic. However, in general, motion blur due to high-speed objects can be expected to cause problems. In this case, a single pixel sees multiple objects with possibly different depths and it is possible that neither of the objects are imaged in perfect focus during detector translation. In tilted and non-planar DOF applications, fast moving scene points can end up being imaged at multiple image locations. All images of a moving scene point would be in-focus if its corresponding 3D positions lie within the (planar/non-planar) DOF. These multiple image locations can be used to measure the velocity and pose of the scene point, as was shown by [21].

Using Different Actuators: In our prototype, we have used a simple linear actuator whose action was synchronized with the exposure time of the detector. However, other more sophisticated actuators can be used. As mentioned above, faster actuators like piezoelectric stacks can dramatically reduce the time needed to translate a detector over the desired distance and so enable low exposure times. This can be very useful for realizing tilted and non-planar DOFs, which need low exposure times. In an EDOF camera, an alternative to a linear actuator is a vibratory actuator – the actuator causes the detector to vibrate with an amplitude that spans the total desired motion of the detector. If the frequency of the vibration is very high (around 100 times within the exposure of an image), then one would not need any synchronization between the detector motion and the exposure time of the detector; errors due to lack of synchronization would be negligible.

Robustness of EDOF Camera PSF: In our experience, the EDOF camera's PSF is robust to the actual motion of the detector or the lens. This is illustrated by the fact, that we are able to capture scenes with large DOFs even when the motion realized is only approximately uniform (example in Section 8). Since this approach does not seem susceptible to small errors in motion, it is particularly attractive for practical implementation in cameras.

Realizing Arbitrary DOFs: We have shown how we can exploit rolling shutter detectors to realize tilted and non-planar DOFs (Sections 6 and 7). In these detectors if the exposure time is sufficiently small, then we can approximately say that the different rows of the image are exposed independently. This allows us to realize DOFs where the focal surfaces are swept surfaces. It is conceivable, that in the future we might have detectors that provide pixel level control of exposure – we can independently control the start and end time of the exposure of each pixel. Such control coupled with a suitable detector motion would enable us to independently choose the scene depth that is imaged in-focus at every pixel, yielding arbitrary DOF manifolds.

Practical Implementation: All DOF manipulations shown in this paper can be realized by moving the lens during image integration (Section 8 shows one example). Compared to moving the detector, moving the lens would be more attractive for camera manufacturers, since cameras already have actuators that move the lens for focusing. All that is needed is to expose the detector while the focusing mechanism sweeps the focal plane through the scene. Hence, implementing these DOF manipulations would not be difficult and can possibly be realized by simply updating the camera firmware.

We believe that flexible DOF cameras can open up a new creative dimension in photography and lead to new capabilities in scientific imaging, computer vision, and computer graphics. Our approach provides a simple means to realizing such flexibility.

ACKNOWLEDGMENTS

The authors would like to acknowledge grants from the National Science Foundation (IIS-04-12759) and the Office of Naval Research (N00014-08-1-0329 and N00014-06-1-0032) that supported parts of this work. Thanks also to Marc Levoy for his comments related to the application of Hausler's method [1] to microscopy.

REFERENCES

[1] G. Hausler, "A Method to Increase the Depth of Focus by Two Step Image Processing," *Optics Communications*, pp. 38–42, 1972.

[2] H. Merklinger, "Focusing the View Camera," 1996.

[3] A. Krishnan and N. Ahuja, "Range estimation from focus using a non-frontal imaging camera," *International Journal of Computer Vision*, pp. 169–185, 1996.

[4] T. Scheimpflug, "Improved Method and Apparatus for the Systematic Alteration or Distortion of Plane Pictures and Images by Means of Lenses and Mirrors for Photography and for other purposes," *GB Patent*, 1904.

[5] H. Nagahara, S. Kuthirummal, C. Zhou, and S. K.Nayar, "Flexible Depth of Field Photography," *European Conference on Computer Vision*, pp. 60–73, 2008.

[6] E. R. Dowski and W. Cathey, "Extended Depth of Field Through Wavefront Coding," *Applied Optics*, pp. 1859–1866, 1995.

[7] N. George and W. Chi, "Extended depth of field using a logarithmic asphere," *Journal of Optics A: Pure and Applied Optics*, pp. 157–163, 2003.

[8] A. Castro and J. Ojeda-Castaneda, "Asymmetric Phase Masks for Extended Depth of Field," *Applied Optics*, pp. 3474–3479, 2004.

[9] A. Levin, R. Fergus, F. Durand, and B. Freeman, "Image and depth from a conventional camera with a coded aperture," *SIGGRAPH*, 2007.

[10] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin, "Dappled photography: mask enhanced cameras for heterodyned light fields and coded aperture," *SIGGRAPH*, 2007.

[11] E. Adelson and J. Wang, "Single lens stereo with a plenoptic camera," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 99–106, 1992.

[12] R. Ng, M. Levoy, M. Brdif, G. Duval, M. Horowitz, and P. Hanrahan, "Light field photography with a hand-held plenoptic camera," *Technical Report Stanford University*, 2005.

[13] T. Georgiev, C. Zheng, B. Curless, D. Salesin, S. K. Nayar, and C. Intwala, "Spatio-angular resolution tradeoff in integral photography," *Eurographics Symposium on Rendering*, pp. 263–272, 2006.

[14] T. Darrell and K. Wohn, "Pyramid based depth from focus," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 504–509, 1988.

[15] S. K. Nayar, "Shape from Focus System," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 302–308, 1992.

[16] M. Subbarao and T. Choi, "Accurate Recovery of Three-Dimensional Shape from Image Focus," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 266–274, 1995.

[17] S. W. Hasinoff and K. N. Kutulakos, "Light-Efficient Photography," *European Conference on Computer Vision*, pp. 45–59, 2008.

[18] A. Agarwala, M. Dontcheva, M. Agrawala, S. Drucker, A. Colburn, B. Curless, D. Salesin, and M. Cohen, "Interactive Digital Photomontage," *SIGGRAPH*, pp. 294–302, 2004.

[19] A. Levin, P. Sand, T. S. Cho, F. Durand, and W. T. Freeman, "Motion-Invarient Photography," *SIGGRAPH*, 2008.

[20] M. Ben-Ezra, A. Zomet, and S. Nayar, "Jitter Camera: High Resolution Video from a Low Resolution Detector," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 135–142, 2004.

[21] O. Ait-Aider, N. Andreff, J.-M. Lavest, and P. Martinet, "Simultaneous Object Pose and Velocity Computation Using a Single View from a Rolling Shutter Camera," *European Conference on Computer Vision*, pp. 56–68, 2006.

[22] D. Field, "Relations between the statistics of natural images and the response properties of cortical cells," *Journal of the Optical Society of America*, pp. 2379–2394, 1987.

[23] H. Hopkins, "The Frequency Response of a Defocused Optical System," *Royal Society of London. Series A, Mathematical and Physical Sciences*, pp. 91–103, 1955.

[24] P. A. Jansson, *Deconvolution of Images and Spectra*. Academic Press, 1997.

[25] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image restoration by sparse 3D transform-domain collaborative filtering," *SPIE Electronic Imaging*, 2008.

[26] L. Yuan, J. Sun, L. Quan, and H.-Y. Shum, "Progressive Inter-scale and intra-scale Non-blind Image Deconvolution," *SIGGRAPH*, 2008.

[27] www.cs.columbia.edu / CAVE / projects / flexible_dof.

[28] P. Burt and R. Kolczynski, "Enhanced image capture through fusion," *International Conference on Computer Vision*, pp. 173–182, 1993.

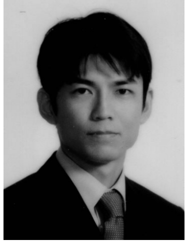
[29] P. Haeberli, "Grafica Obscura," *www.sgi.com/grafica/*, 1994.

[30] Q. Shan, J. Jia, and A. Agarwala, "High-quality Motion Deblurring from a Single Image," *SIGGRAPH*, 2008.

[31] N. Asada, H. Fujiwara, and T. Matsuyama, "Seeing behind the scene: Analysis of photometric properties of occluding edges by the reversed projection blurring model," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 155–167, 1998.

[32] S. Bhasin and S. Chaudhuri, "Depth from defocus in presence of partial self occlusion," *International Conference on Computer Vision*, pp. 488–493, 2001.

Sujit Kuthirummal received his PhD degree in Computer Science from Columbia University in 2009. Prior to that, he received his B. Tech and M.S. degrees in Computer Science from the International Institute of Information Technology, Hyderabad in 2002 and 2003, respectively. Since 2009 he has been a Member of Technical Staff at Sarnoff Corporation. His research interests include computational photography and 3D reconstruction.



Hajime Nagahara is an Assistant Professor at the Graduate School of Engineering Science, Osaka University, Japan. He received B. E. and M. E. degrees in Electrical and Electronic Engineering from Yamaguchi University in 1996 and 1998, respectively. He received his Ph.D. in System Engineering from Osaka University in 2001. He was a Research Associate of the Japan Society for the Promotion of Science between 2001-2003. He was a Research Associate of the Graduate School of Engineering Science, Osaka University, between 2003-2006. He was a Visiting Associate Professor at CREA University of Picardie Jules Verns, France in 2005. He was a Visiting researcher at Columbia University, USA in 2007-2008. Image processing, computer vision and virtual reality are his research fields. He received an ACM VRST2003 Honorable Mention Award in 2003.

Changyin Zhou received the BS degree in Statistics and the MS degree in Computer Science from Fudan University in 2001 and 2007, respectively. He is currently a doctoral student in the Computer Science Department of Columbia University. His research interests include computational imaging and physics-based vision. He is a student member of the IEEE.

Shree K. Nayar received his PhD degree in Electrical and Computer Engineering from the Robotics Institute at Carnegie Mellon University in 1990. He is currently the T. C. Chang Professor of Computer Science at Columbia University. He co-directs the Columbia Vision and Graphics Center. He also heads the Columbia Computer Vision Laboratory (CAVE), which is dedicated to the development of advanced computer vision systems. His research is focused on three areas; the creation of novel cameras, the design of physics based models for vision, and the development of algorithms for scene understanding. His work is motivated by applications in the fields of digital imaging, computer graphics, and robotics.

He has received best paper awards at ICCV 1990, ICPR 1994, CVPR 1994, ICCV 1995, CVPR 2000 and CVPR 2004. He is the recipient of the David Marr Prize (1990 and 1995), the David and Lucile Packard Fellowship (1992), the National Young Investigator Award (1993), the NTT Distinguished Scientific Achievement Award (1994), the Keck Foundation Award for Excellence in Teaching (1995), the Columbia Great Teacher Award (2006) and the Carnegie Mellon University Alumni Achievement Award (2009). In February 2008, he was elected to the National Academy of Engineering.