

Flexible Depth of Field Photography^{*}

Hajime Nagahara^{1,2}, Sujit Kuthirummal²,
Changyin Zhou², and Shree K. Nayar²

¹ Osaka University

² Columbia University

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 three 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. Applying deconvolution to a captured image gives an image with extended DOF and yet 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. Finally, we show that our camera can capture images with tilted DOFs (Scheimpflug imaging) without tilting the image detector. We believe flexible DOF imaging can open a new creative dimension in photography and lead to new capabilities in scientific imaging, vision, and graphics.

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

^{*} Parts of this work were supported by grants from the National Science Foundation (IIS-04-12759) and the Office of Naval Research (N00014-08-1-0329 and N00014-06-1-0032.).

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, irrespective of the complexity of the scene, the captured image can be deconvolved with a single, known blur kernel to recover an image with significantly greater DOF. 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 the small distances involved in microscopy.
- **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 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. Furthermore, by translating the image detector *non-uniformly* (varying speed), we can emulate a non-planar image detector. This allows us to focus on curved surfaces in the scene.

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. We believe our approach can open up a new creative dimension in photography and lead to new capabilities in scientific imaging, computer vision, and computer graphics.

2 Related Work

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 [5,6,7]. Thus, by deconvolving the captured image with a single blur kernel, one can obtain an all-focused image. In this case, 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. Our approach has greater flexibility as it can even be used to achieve discontinuous or tilted DOFs.

Recently, Levin et al. [8] and Veeraraghavan et al. [9] have used masks at the lens aperture to control the properties of the defocus blur kernel. From a single captured photograph, 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 irrespective of scene complexity. Also, the masks used in both 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 [10,11,12]. 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 resolutions when compared to our approach.

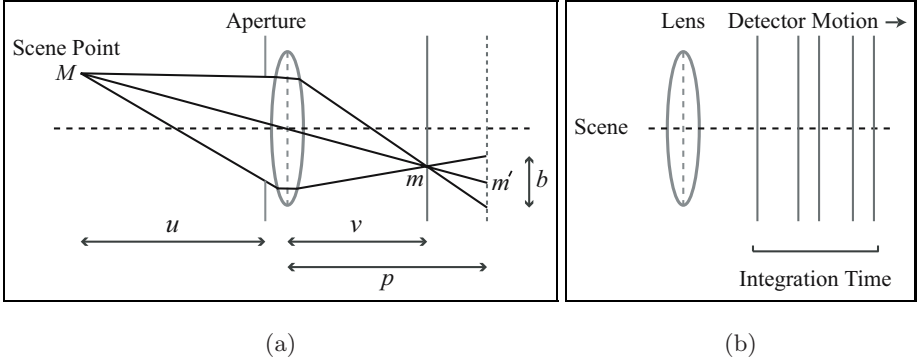


Fig. 1. (a) A scene point M , at a distance u from the lens, is imaged in perfect focus by a detector at a distance v from the lens. If the detector is shifted to a distance p from the lens, M is imaged as a blurred circle with diameter b centered around m' . (b) Our flexible DOF camera translates the 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.

A related approach is to capture many images to form a focal stack [13,14,15]. 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. Our approach does not recover scene depth, but can produce an all-in-focus photograph from a single, well-exposed image.

There is similar parallel work on moving the detector during image integration [16]. 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. [17] 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 [18]. We show how such a detector can be used to focus on tilted scene planes.

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 Gaussian lens law:

$$\frac{1}{f} = \frac{1}{u} + \frac{1}{v}. \quad (1)$$

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

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

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:

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

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:

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

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

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

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 2(a) 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 DOFs, we used a 1/2.5" Micron CMOS detector (with 2592x1944 pixels) which has a rolling shutter.

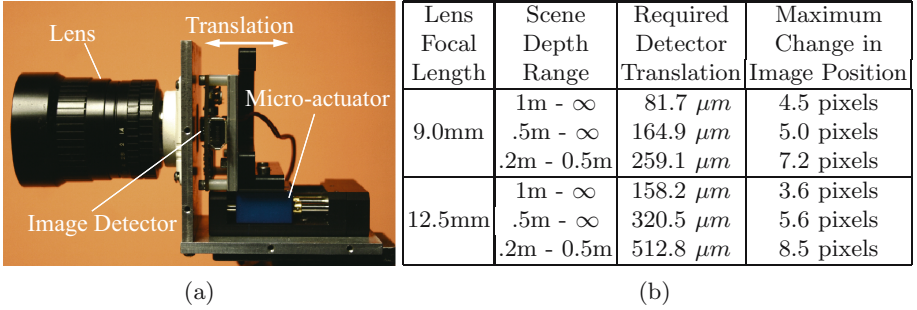


Fig. 2. (a) Prototype system with flexible DOF. (b) Translation of the detector required for sweeping the focal plane through different scene depth ranges. The maximum change in the image position of a scene point that results from this translation, when a 1024x768 pixel detector is used, is also shown.

The table in Figure 2(b) 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. The fourth column of the table in Figure 2(b) 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 perspective 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 the 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 that the PSF of the lens can be modeled using the pillbox function in Equation 3, the IPSF in Equation 5 simplifies to

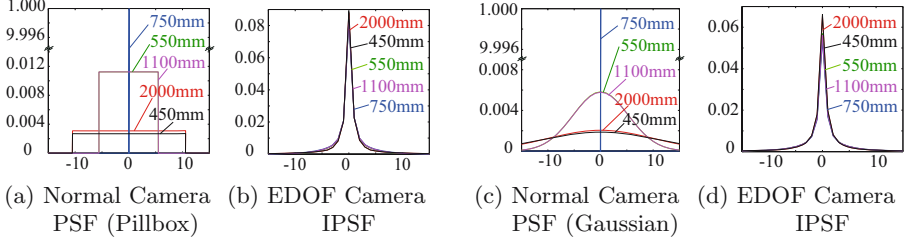


Fig. 3. Simulated (a,c) normal camera PSFs and (b,d) EDOF camera IPSFs, obtained using pillbox and Gaussian lens PSF models for 5 scene depths. Note that the IPSFs are almost invariant to scene depth.

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

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

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

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 $T = 360$ msec, 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*.

To verify this empirical observation, 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 $T = 360$ msec. The detector motion parameters for the EDOF camera are $p(0) = 12.5$ mm and $s = 1$ mm/sec. The first column of Figure 4 shows the measured PSF at the center pixel of the normal camera for 5 different 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 spatial location and scene depth. This also validates our claim that the small magnification changes that arise due to detector motion (discussed in Section 3) do not have a significant impact on the IPSFs.

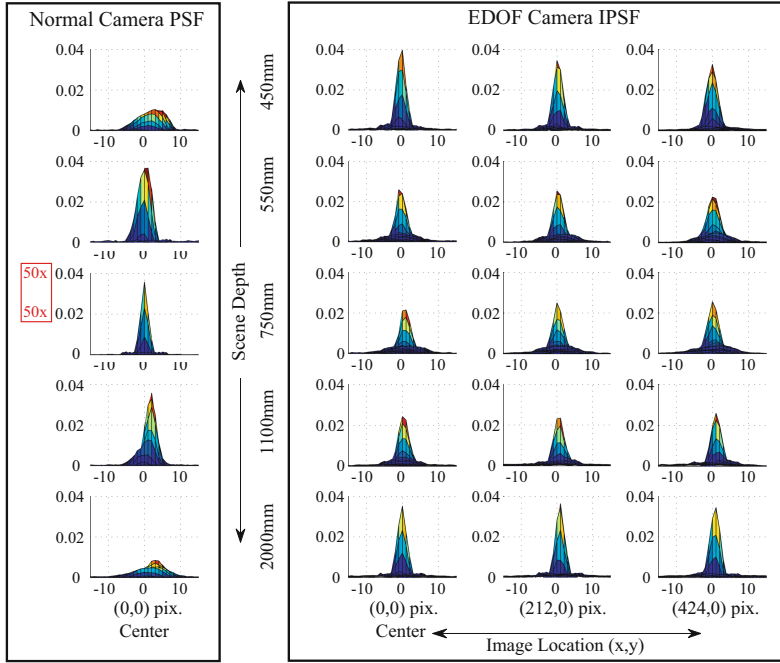


Fig. 4. (Left column) The measured PSF of a normal camera shown for 5 different scene depths. Note that the scale of the plot in the center row is 50 times that of the other plots. (Right columns) The measured IPSF of our EDOF camera shown for different scene depths (vertical axis) and image locations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

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 [19] being two popular ones. For our results, we have used the approach of Dabov et al. [20], 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.

Figure 5(a) shows an image captured by our EDOF camera. It was captured with a 12.5 mm Fujinon lens with $f/1.4$ and 0.36 second exposure. Notice that the captured image looks slightly blurry, but high frequencies of all scene elements are captured. This scene spans a depth range of approximately 450 mm to 2000 mm – 10 times larger than the DOF of a normal camera with identical lens settings. Figure 5(b) shows the EDOF image computed from the captured image, in which the entire scene appears focused¹. Figure 5(c) shows the image

¹ Mild ringing artifacts in the computed EDOF images are due to deconvolution.

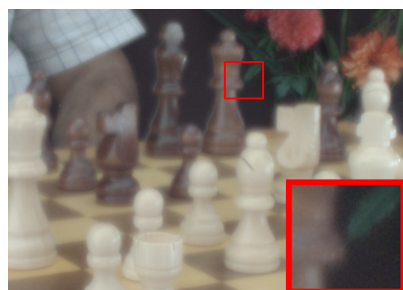
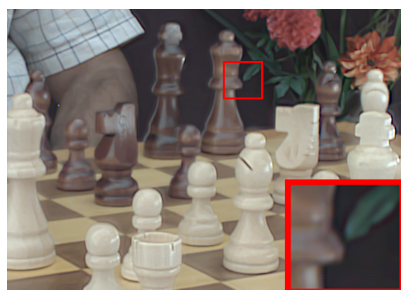
captured by a normal camera with identical $f/\#$ and exposure time. The nearest scene elements are in focus, while the farther scene elements are severely blurred. The image captured by a normal camera with the same exposure time, but with a smaller aperture ($f/8$) is shown in Figure 5(d). The intensities of this image were scaled up so that its dynamic range matches that of the corresponding computed EDOF image. All scene elements look reasonably sharp, but the image is very noisy as can be seen in the inset (zoomed). The computed EDOF image has much less noise, while having comparable sharpness. Figures 5(e-h) show 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. In short, our EDOF camera can capture scenes with large DOFs as well as high SNR. High resolution versions of these images as well as other examples can be seen at [21].

4.3 Analysis of SNR Benefits of EDOF Camera

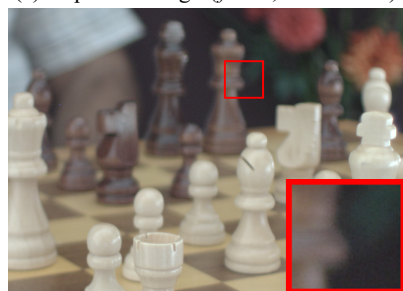
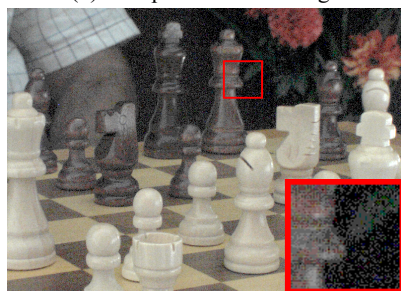
Deconvolution using Dabov et al.'s method [20] produces visually appealing results, but since it has a non-linear denoising step, it is not suitable for analyzing the SNR of deconvolved captured images. Therefore, we performed a simulation that uses Wiener deconvolution [19]. Given an IPSF k , we convolve 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 6(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 other parameters) as in our EDOF experiments (Section 4.2). 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 6(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 of the table shows the effective DOFs realized; the normal camera is assumed to be focused at a scene distance that corresponds to the center position of the detector motion. One can see that, as the noise level in the captured image increases, the SNR benefits of EDOF cameras increase. 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 with $f/2.8$, but produces the DOF of a normal camera with $f/8$.

(a) Captured Image ($f/1.4$, $T=0.36\text{sec}$)

(b) Computed EDOF Image

(c) Image from Normal Camera
($f/1.4$, $T=0.36\text{sec}$, Near Focus)(d) Image from Normal Camera
($f/8$, $T=0.36\text{sec}$, Near Focus) with Scaling(e) Captured Image ($f/1.4$, $T=0.72\text{sec}$)

(f) Computed EDOF Image

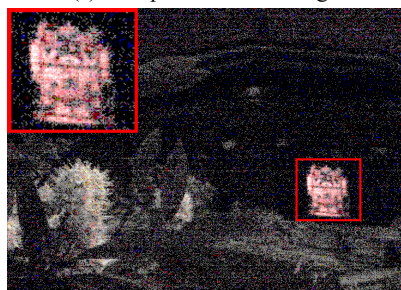
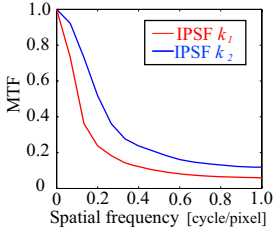
(g) Image from Normal Camera
($f/1.4$, $T=0.72\text{sec}$, Near Focus)(h) Image from Normal Camera
($f/8$, $T=0.72\text{sec}$, Near Focus) with Scaling

Fig. 5. (a,e) Images captured by the EDOF camera. (b,f) EDOF images computed from images in (a) and (e), respectively. Note that the entire scene appears focused. (c,g) Images captured by a normal camera with identical settings, with the nearest object in focus. (d,h) Images captured by a normal camera at $f/8$.



(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. 6. (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 given.

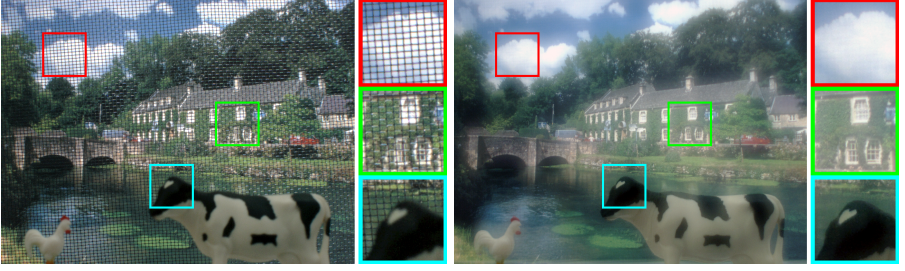
In the above analysis, the SNR was averaged over all frequencies. However, it must be noted that SNR is frequency dependent - SNR is greater for lower frequencies than for higher frequencies in the deconvolved EDOF images. Hence, high frequencies in an EDOF image would be degraded, compared to the high frequencies in a perfectly focused image. However, in our experiments this degradation is not strong, as can be seen in the full resolution images at [21].

5 Discontinuous Depth of Field

Consider the image in Figure 7(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, 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 7(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.

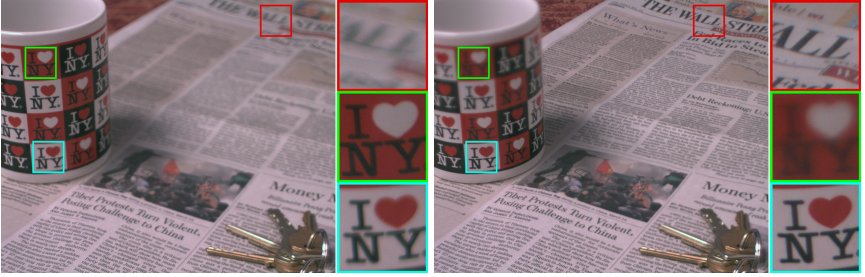
6 Tilted Depth of Field

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



(a) Image from Normal Camera ($f/11$) (b) Image from Our Camera ($f/1.4$)

Fig. 7. (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 ($f/1.4$, $T=0.03\text{sec}$) (b) Image from our Camera ($f/1.4$, $T=0.03\text{sec}$)

Fig. 8. (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.

DOF camera can be programmed to focus on tilted scene planes by simply translating (as in the 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. 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 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]:

$$\theta = \tan^{-1}\left(\frac{sT}{H}\right) \quad \text{and,} \quad \phi = \tan^{-1}\left(\frac{2f \tan(\theta)}{2p(0) + H \tan(\theta) - 2f}\right). \quad (8)$$

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 8 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 from Figure 8(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 8) and capture the table top (with the newspaper and keys) in focus, as shown in Figure 8(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. It is interesting to note that, by translating the detector with varying speed, we can emulate non-planar detectors, that can focus on curved scene surfaces.

7 Discussion

In this paper we have proposed a camera with a flexible DOF. DOF is manipulated in various ways by changing the position of the detector during image integration. We have shown how such a system can capture arbitrarily complex scenes with extended DOF and high SNR. 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. All of these functionalities are achieved by simply controlling the motion of the detector during the exposure of a single image.

While computing images with extended DOF, we have not explicitly modeled occlusions at depth discontinuities or motion blur caused by object/camera motion. Due to defocus blur, images 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. With respect to motion blur, we have not observed any visible artifacts in EDOF images computed for scenes with typical object motion (see Figure 5). However, 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. It is possible that neither of the objects are imaged in perfect focus during detector translation. This scenario is an interesting one that warrants further study.

In addition to the DOF manipulations shown in this paper, we have (a) captured extended DOF video by moving the detector forward one frame, backward the next, and so on (the IPSF is invariant to the direction of motion), (b) captured scenes with non-planar DOFs, and (c) exploited the camera's focusing

mechanism to capture extended DOF by manually rotating a SLR camera lens' focus ring during image integration. For lack of space, we have not included these results here; they can be seen at [21].

References

1. Hausler, G.: A Method to Increase the Depth of Focus by Two Step Image Processing. *Optics Communications*, 38–42 (1972)
2. Merklinger, H.: *Focusing the View Camera* (1996)
3. Krishnan, A., Ahuja, N.: Range estimation from focus using a non-frontal imaging camera. *IJCV*, 169–185 (1996)
4. Scheimpflug, T.: 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. Dowski, E.R., Cathey, W.T.: Extended Depth of Field Through Wavefront Coding. *Applied Optics*, 1859–1866 (1995)
6. George, N., Chi, W.: Extended depth of field using a logarithmic asphere. *Journal of Optics A: Pure and Applied Optics*, 157–163 (2003)
7. Castro, A., Ojeda-Castaneda, J.: Asymmetric Phase Masks for Extended Depth of Field. *Applied Optics*, 3474–3479 (2004)
8. Levin, A., Fergus, R., Durand, F., Freeman, B.: Image and depth from a conventional camera with a coded aperture. *SIGGRAPH* (2007)
9. Veeraraghavan, A., Raskar, R., Agrawal, A., Mohan, A., Tumblin, J.: Dappled photography: mask enhanced cameras for heterodyned light fields and coded aperture. *SIGGRAPH* (2007)
10. Adelson, E., Wang, J.: Single lens stereo with a plenoptic camera. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 99–106 (1992)
11. Ng, R., Levoy, M., Brdif, M., Duval, G., Horowitz, M., Hanrahan, P.: Light field photography with a hand-held plenoptic camera. Technical Report Stanford University (2005)
12. Georgiev, T., Zheng, C., Curless, B., Salesin, D., Nayar, S.K., Intwala, C.: Spatio-angular resolution tradeoff in integral photography. In: *Eurographics Symposium on Rendering*, pp. 263–272 (2006)
13. Darrell, T., Wohn, K.: Pyramid based depth from focus. *CVPR*, 504–509 (1988)
14. Nayar, S.K.: Shape from Focus System. *CVPR*, 302–308 (1992)
15. Subbarao, M., Choi, T.: Accurate Recovery of Three-Dimensional Shape from Image Focus. *PAMI*, 266–274 (1995)
16. Levin, A., Sand, P., Cho, T.S., Durand, F., Freeman, W.T.: Motion-Invariant Photography. *SIGGRAPH*, *ACM Transaction on Graphics* (2008)
17. Ben-Ezra, M., Zomet, A., Nayar, S.: Jitter Camera: High Resolution Video from a Low Resolution Detector. *CVPR*, 135–142 (2004)
18. Ait-Aider, O., Andreff, N., Lavest, J.M., Martinet, P.: Simultaneous Object Pose and Velocity Computation Using a Single View from a Rolling Shutter Camera. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) *ECCV 2006*. LNCS, vol. 3952, pp. 56–68. Springer, Heidelberg (2006)
19. Jansson, P.A.: *Deconvolution of Images and Spectra*. Academic Press, London (1997)
20. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: Image restoration by sparse 3D transform-domain collaborative filtering. *SPIE Electronic Imaging* (2008)
21. www.cs.columbia.edu/CAVE/projects/flexible_dof