IPSJ Transactions on Computer Vision and Applications Vol.4 1–11 (Mar. 2012)

[DOI: 10.2197/ipsjtcva.4.1]

Research Paper

Programmable Aperture Camera Using LCoS

Hajime Nagahara[1,a)] Changyin Zhou[2] Takuya Watanabe[3]
Hiroshi Ishiguro[3] Shreee K. Nayar[2]

Received: April 28, 2011, Accepted: December 20, 2011, Released: March 8, 2012

Abstract: Since 1960s, aperture patterns have been studied extensively and a variety of coded apertures have been proposed for various applications, including extended depth of field, defocus deblurring, depth from defocus, light field acquisition, etc. Researches have shown that optimal aperture patterns can be quite different due to different applications, imaging conditions, or scene contents. In addition, many coded aperture techniques require aperture patterns to be temporally changed during capturing. As a result, it is often necessary to have a *programmable aperture camera* whose aperture pattern can be dynamically changed as needed in order to capture more useful information. In this paper, we propose a programmable aperture camera using a Liquid Crystal on Silicon (LCoS) device. This design affords a high brightness contrast and high resolution aperture with a relatively low light loss, and enables one change the pattern at a reasonably high frame rate. We build a prototype camera and evaluate its features and drawbacks comprehensively by experiments. We also demonstrate three coded aperture applications in defocus deblurring, depth from defocus and light field acquisition.

Keywords: coded aperture, computational photography, deblurring, depth from defocus, light field acquisition, LCoS

1. Introduction

In the past decades, coded aperture techniques have been studied extensively in optics, computer vision and computer graphics, and a variety of coded aperture techniques have been proposed for various applications. The optimal aperture patterns can be quite different from one application to another. For defocus deblurring, coded apertures are optimized to be broad-band in the Fourier domain [1], [2]. For depth from defocus, coded apertures are optimized to have more zero-crossing frequencies [3], [4]. For multiplexing light field acquisition, an optimal set of aperture patterns are solved for the best signal-to-noise ratio (SNR) after de-multiplexing [5]. Aperture can also be coded in the temporal dimension for motion deblurring [6]. Coded aperture methods have also been used in many other applications, including lensless imaging [7], [15], natural matting [8], etc. **Figure 1** shows a collection of some coded apertures that were proposed in the past.

There are many situations where the aperture pattern should be dynamically updated as needed. First, from the aspect of information capturing, ideally aperture pattern should be adaptive to scene contents. For example, the pattern should be optimized for defocus deblurring if the scene has a large depth, and it should be optimized for motion deblurring if the scene has many objects in motion. Secondly, aperture pattern should be adaptive to the specific application purpose. For example, people have shown that a coded aperture optimized for defocus deblurring is often a bad choice for depth from defocus [4], and multiplexing light field acquisition technique requires different aperture codings for different target angular resolutions. Thirdly, the pattern should be adaptive to the imaging condition. For example, the optimal aperture pattern for defocus deblurring is different at different image noise levels as shown in Zhou's work [2]. In addition, some coded aperture techniques need to capture multiple images with different aperture patterns (e.g., coded exposure [6], coded aperture pair [4] and multiplexed light field acquisition [5]). In all these situations, people need a *programmable aperture camera* whose aperture pattern can be updated at a reasonable speed.

In literature, people has used transmissive liquid crystal displays (LCD) to control aperture patterns [5], [15]. However, the LCD implementation has severe drawbacks. The electronic elements on LCD pixels occlude lights and lead to a low light efficiency. These occluders also cause strong and complicated defocus and diffraction artifacts. These artifacts can be very strong and eliminate the benefits of aperture codings. Consider the popular applications of coded aperture (e.g., defocus deblurring, depth from defocus), we argue that a good programmable aperture is necessary to have the following features:

(1) Easy mount. For different applications or scenes, people may use different lenses and sensors. Therefore, it is important to build a programmable aperture that can be easily mounted to different lenses and sensors.

(2) High light efficiency. The loss of light leads to decreased SNR. As shown in some papers [2], [9], a high light efficiency is the key to achieve high performance in defocus deblurring, depth from defocus, multiplexing light field acquisition, etc.

(3) Reasonable frame rate. Some coded aperture techniques

[1] Kyushu University, Fukuoka 819–0395, Japan
[2] Columbia University, New York, USA
[3] Osaka University, Toyonaka, Osaka 560–8531, Japan
[a)] nagahara@ait.kyushu-u.ac.jp

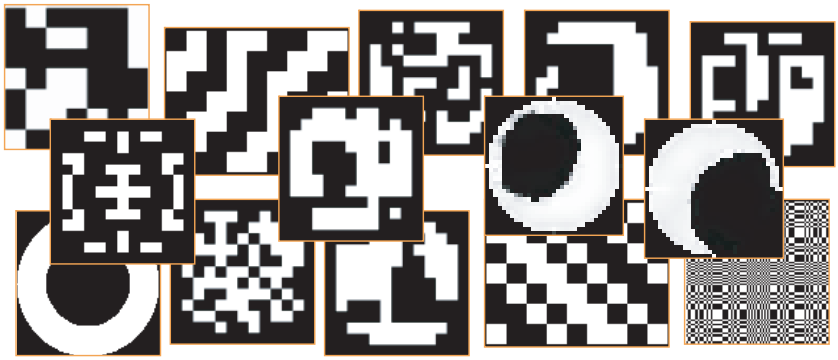
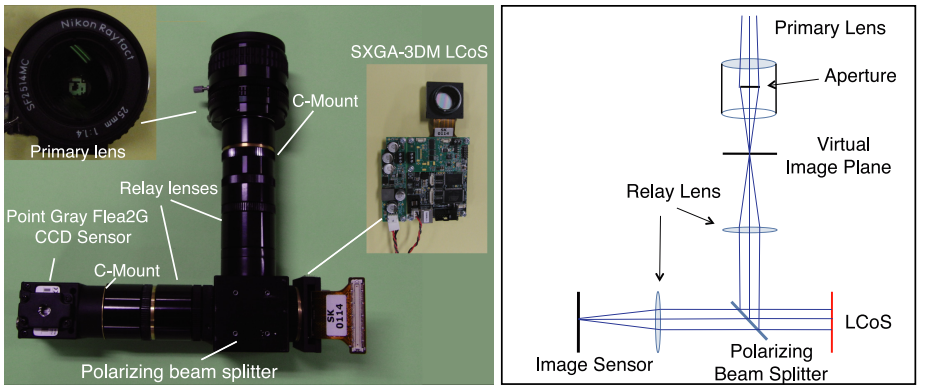


Fig. 1 A variety of coded aperture patterns proposed for various applications.



(a) Our prototype programmable aperture camera (b) The optical diagram of the prototype camera

Fig. 2 Programmable aperture camera using an LCoS device. (a) Our prototype LCoS programmable aperture camera. In the left-top corner is the Nikon F/1.4 25 mm C-mount lens that is used in our experiments. On the right is an LCoS device. (b) The optical diagram of the proposed LCoS programmable aperture camera.

capture multiple images of a scene using different aperture patterns [4], [5]. For dynamic scenes, these techniques require multiple images to be captured within a reasonable short time in order to reduce motion blur, and at the same time, the aperture pattern must also be updated at the same frame rate and be synchronized with the sensor exposure.

(4) High brightness contrast. Most optimized aperture patterns in literature have high brightness contrast - many of them are binary patterns. We may fail to display optimized patterns without a high brightness contrast.

To meet these requirements, we propose in this paper a programmable aperture camera by using a Liquid Crystal on Silicon (LCoS) device as shown in **Fig. 2**. LCoS is a reflective liquid crystal device that has a high fill factor (92%) and high reflectivity (60%). Compared with transmissive LCD, an LCoS device usually suffers much less from light loss and diffraction. Figure 2 shows the structure of our proposed programmable aperture camera. The use of LCoS device in our prototype camera enables us to dynamically change aperture patterns as needed at a high resolution (1280×1024 pixels), a high frame rate (5 kHz maximum), and a high brightness contrast. By using the relay optics, we can mount any C-Mount or Nikon F-Mount lens to our programmable aperture camera. Remarkably, our implementation used only off-the-shelf elements and people may reproduce or even improve the design for their own applications.

A detailed description and analysis to our proposed system will be given in Section 3. The features and limitations of the present prototype camera are evaluated via experiments in Section 4. The proposed coded aperture camera can be a platform to implement many coded aperture techniques. As examples, in Section 5, we demonstrate the use of our prototype camera in three applications: defocus deblurring [1], [2], depth from defocus [4] and multiplexing light field acquisition [5]. This paper is the extended version of the paper [25] that appeared in ECCV2010.

2. Related Work

Coded aperture technique was first introduced in the field of high energy astronomy in 1960s as a novel way of improving SNR for lensless imaging of x-ray and γ-ray sources [10]. It is also in the 1960s that researchers in optics began developing unconventional apertures to capture high frequencies with less attenuation. In the following decades, many different aperture patterns were proposed (e.g., apodizations [11], [12], [13], [14], and MURA [7]).

Coded aperture research resurfaces in computer vision and graphics in recent years. People optimize coded aperture patterns to be broad-band in the Fourier domain in order that more information can be preserved during defocus for the later deblurring [1], [2]. Levin et al. [3] optimizes a single coded aperture to have more zero-crossings in the Fourier domain so that the depth information can be better encoded in a defocused image. Zhou et al. [4] show that by using the optimized coded aperture pair,

they will be able to simultaneously recover a high quality focused image and a high quality depth map from a pair of defocused images. In Ref. [5], Liang et al. proposed to take a bunch of images using different coded aperture patterns in order to capture the light field.

Coded apertures have also been used for many other applications. Zomet and Nayar propose a lensless imaging technique by using an LCD aperture [15]. Raskar et al. use a coded flutter shutter aperture for motion deblurring [6].

Coded aperture camera can be implemented in several ways. One popular coded aperture implementation is to disassemble the lens and insert a mask, which can be made of a printed film or even a cut paper board [1], [2], [3]. The major disadvantages of this method are that one has to disassemble the lens, and that the pattern cannot be easily changed once the mask is inserted. Note that most commercial lenses cannot be easily disassembled without serious damages. People have also used some mechanical ways to modify apertures. Aggarwal and Ahuja propose to split the aperture by using a half mirror for high dynamic range imaging [16]. Green et al. build a complicated mechanical system and relay optics to split a circular aperture into three parts of different shapes [17]. Dowski et al. proposed wave-front coding [18] which places special optical element, called cubic phase plate, on an aperture position for extended depth from field by deblurring.

To dynamically change aperture patterns during capturing, people have proposed to use transmissive liquid crystal display (LCD) devices as in the works [5], [15]. One problem with the LCD implementation is that the electronic elements sit in the LCD pixels not only block a significant portion of incoming light but also cause significant diffractions. Some custom LCDs are designed to have a higher light efficiency. However, these LCDs usually either have much low resolution (e.g., 5×5 pixels in Liang's implementation [5]) or are prohibitively expensive. In this paper, we propose to use a reflective liquid crystal on silicon (LCoS) device [19], which has much higher light efficiency and suffers less from diffraction. LCoS has been used before in computer vision for high dynamic range imaging [20]. Another similar device that could be used to modulate apertures is the digital micro-mirror device (DMD). Nayar and Branzoi use a DMD device to control the irradiance to each sensor pixel for various applications, including high dynamic range and feature detection [21]. However, each DMD pixel only has two states and therefore DMD devices can only be used to implement binary patterns. More importantly, existing DMD is restricted mirror angles to ±12 degrees only (DMD cannot take 0 degree as a mirror angle). If we apply DMD to an aperture in same fashion with our LCoS implementation, the aperture is slanted to an optical axis and coding result is optically different from regular coded aperture because of the restriction.

3. Optical Design and Implementation

We propose to implement a programmable aperture camera by using a liquid crystal on silicon (LCoS) device as an aperture. LCoS is a reflective micro-display technique typically used in projection televisions. An LCoS device can change the polarization direction of rays that are reflected by each pixel. Compared with the typical transmissive LCD technique, it usually produces higher brightness contrast and higher resolution. Furthermore, LCoS suffers much less from light loss and diffraction than LCD does. This is because the electronic components sitting on each pixel of LCD device block lights and cause significant diffraction, and on the contrary, an LCoS device has all the electronic components behind the reflective surface and therefore provides much higher fill factors.

Table 1 Specification of LCoS device.

Resolution	1280×1024 pixels
Reflective depth	8 bits
Pixel fill factor	>92%
Reflectivity	60%
Contrast ratio	400:1
Physical dimension	17.43×13.95 mm
Switching pattern	40 μs

One of our major design goals is to make the primary lens separable from the programmable aperture in order that people can directly attach any compatible lenses without disassembling the lens. To achieve this, we propose to integrate an LCoS device into relay optics.

As shown in Fig. 2, our proposed system consists of a primary lens, two relay lenses, one polarizing beam splitter, an LCoS device, and an image sensor. Only off-the-shelf elements are used in our prototype camera implementation. We choose a Forth dimension display SXGA-3DM LCoS micro-display. **Table 1** shows the specifications of this LCoS device. We use two aspherical lenses of 100 mm and 125 mm focal lengths (Edmund Optics, part #447641 and #47642) for each lens of the relay optics. The compound focal length of the lens is 55 mm. We choose a cubic polarizing beam splitter (Edmund Optics, part #49002), and a Point Grey Grasshopper GRAS-14S5C-C camera (2/3″ CCD, 1384×1036 pixels at 25 fps). The camera shutter is synchronized with the LCoS device by using an output trigger (25 Hz [*1]) of the LCoS driver.

People have a plenty of freedom in choosing primary lenses for this system. The primary lens and the image sensor are attached to the optics via the standard C-mount. Therefore, a variety of C-mount cameras and lenses can be directly used with this prototype system. SLR lenses (e.g., Nikon F-mount lenses) can also be used via a proper lens adopter. In our experiments, we use a Nikon Rayfact 25 mm F/1.4 C-mount lens.

We can see from Fig. 2 (b) that an incoming light from a scene is first collected by the primary lens and focused at the virtual image plane. A cone of light from each pixel of the virtual image plane is then forwarded by the first relay lens to the polarizing beam splitter. The beam splitter separates the light into S-polarized and P-polarized (linear polarizations perpendicular to each other) lights by reflection and transmission, respectively. The reflected S-polarized light is further reflected by LCoS. The LCoS device can rotate the polarization direction at every pixel by arbitrary degrees. For example, if the pixel on LCoS is set to 255 (8 bit depth), the polarization of the light is rotated by 90 degree and becomes P-polarized, and then the light will pass through the

[*1] Note that the LCoS can be modulated at 5 kHz maximum. We use 25 Hz in order that it can be synchronized with the sensor.

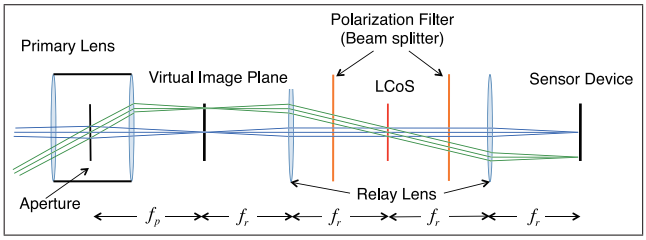


Fig. 3 An equivalent optical diagram to that in Fig. 2 (b). The virtual image plane and the sensor plane are conjugated by the relay lens. The LCoS is the aperture stop of the system.

splitter and reach to the sensor. If the pixel on LCoS is set to 0, the polarization will not be changed by LCoS and the reflected light will be blocked by the splitter.

Consider the LCoS device as a mirror, the diagram in Fig. 2 (b) can be easily shown equivalent to that in **Fig. 3**. The proposed optics can be better understood from Fig. 3. The virtual image plane is placed f_p from the aperture position of primary lens. The distances from the plane to first relay lens, the lens to LCoS, the LCoS to the second relay lens, and the lens to sensor device are f_r for realizing relay optics. Here, f_p is the focal length of the primary lens and f_r is that of relay lens ($f_p = 25$ mm, $f_r = 55$ mm in our prototype). As a result, the sensor plane is conjugate to the virtual image plane. The LCoS device is relatively smaller than other stops in this optical system and works as the aperture stop.

4. Optical Analysis and Experimental Evaluation

We have analyzed and evaluated the prototype camera for estimating the specification or current limitations.

4.1 Effective F-number

Since the LCoS device works as the aperture stop in the proposed system, F-number ($f/\#$) of the primary lens is no longer the effective $f/\#$ of the camera. The actual $f/\#$ of the system is decided by focal length of the relay lens f_r and physical size of LCoS. For a circular aperture, $f/\#$ is usually defined as the ratio of focal length to the aperture diameter. For the rectangle nature of the LCoS, we use $2\sqrt{uv/\pi}$ as the diameter, where (u, v) is the dimension of LCoS. Therefore have:

$$f/\# = \frac{2}{f_r}\sqrt{\frac{uv}{\pi}}. \tag{1}$$

According to Eq. (1), the effective $f/\#$ of the prototype can be computed as $f/2.96$, while the $f/\#$ of the primary lens is $f/1.4$.

4.2 Field of View

Figure 3 shows that the relay system copies the virtual image to sensor plane by a magnification ratio of $1:1$. Therefore, the field of view (FOV) of the proposed camera is the same as if the sensor were placed at the virtual image plane. The FOV can be estimated by using the sensor size and the effective focal length of the primary lens:

$$\mathrm{FOV} \approx 2\arctan\frac{d}{2f_p}, \tag{2}$$

where d is a diagonal size of the sensor and f_p is the effective focal length of the primary lens.

Our prototype camera uses a 25 mm lens and therefore the camera FOV can be computed as 24.8° according to Eq. (2). Of course, we can change the FOV by using a primary lens with a different focal length.

4.3 Light Efficiency

Light efficiency is one of the most important index in a coded aperture camera. Ideally, the light efficiency of our prototype camera is calculated by:

$$27.6\% = 50\%(\text{polarization}^{*2}) \times 92\%(\text{fillfactor}) \times 60\%(\text{reflectivity}). \tag{3}$$

We notice that many other optical elements in the camera (e.g., a beam splitter, and two relay lenses, etc.) may also attenuate the intensity of captured images. To measure the light efficiency accurately, we captured two images of a uniformly white plane. One image was captured using our prototype camera, and another image was captured without the LCoS aperture (the same sensor and the same lens with $f/\#$ set to 2.8). The ratio of the averaged brightness of these two captured images is computed as 37.85:229.4, which indicates the light efficiency of the system. The light efficiency of our system is about 16.5%.

The theoretical light efficiency of a transmissive LCD[*3] can also be calculated using a similar formula:

$$7.4\% = 50\%(\text{polarization}^{*4}) \times 55\%(\text{fillfactor}) \times 27\%(\text{transmittance}). \tag{4}$$

The light efficiency of our LCoS implementation is at least three times higher than that of the LCD implementation.

4.4 Vignetting

From the two images captured with and without the LCoS aperture, we can compute the vignetting curves of the prototype camera and a normal camera. The horizontal vignetting curves of our prototype camera and a normal camera are shown in **Fig. 4** in red and blue solid lines, respectively. The corresponding dashed lines show the vertical vignetting curves. Figure 4 shows intensity attenuation by vignetting especially on horizontal direction. This is caused by additional relay lens optics. However, once we estimate the effect as shown in Fig. 4 as a look up table, we can easily calibrate it.

4.5 Transmission Fidelity

Another important quality index of a coded aperture implementation is the transmission fidelity – the consistence between the

[*2] A polarized beam splitter splits incoming lights based on their polarizations. Although the light interacts with the splitter twice, the light efficiency of beam splitter is still 50%. This is because 100% light will pass through the splitter at the second interaction when its polarization is aligned to that of the splitter.

[*3] Note that the fill factor or transmittance of the LCD can be slightly different due to different implementations (e.g., physical sizes and resolutions). We assume a typical LCD with a similar physical size and resolution to the LCoS used in our implementation.

[*4] An transmissive LCD also needs polarizers which are placed on both side of the LCD panel. The LCD panel just twist a ray polarization as same as LCoS. The difference between LCoS and LCD is reflecting on or transmitting though liquid crystal material for changing the polarization.

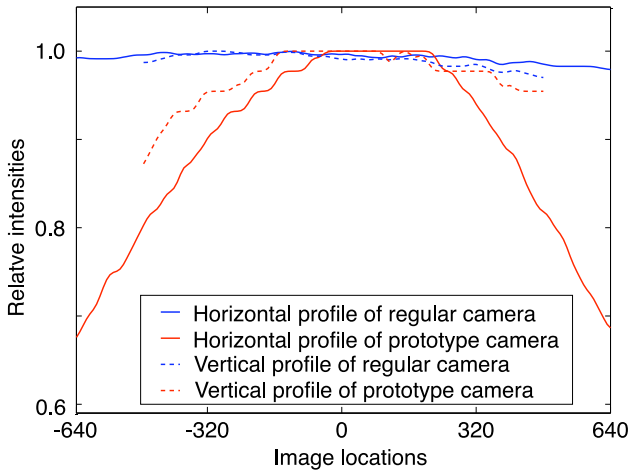


Fig. 4 Vignetting profiles. The red and blue solid lines indicate the horizontal vignetting curves of the prototype camera and a regular camera, respectively. The dashed lines indicate their vertical vignetting profiles.

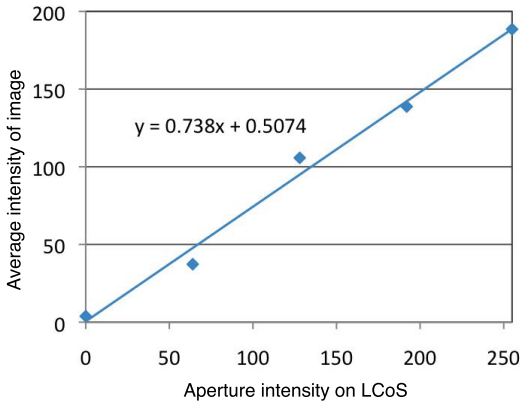


Fig. 5 The aperture transmittance is linear to the LCoS intensity.

actual transmittance of coded aperture and the input intensity of the LCoS device. To evaluate the transmission fidelity, we captured images of uniformly white plane using circular apertures of different intensities. **Figure 5** shows the average intensity of captured images with respect to the input intensities of the circular aperture (implemented using LCoS device). The linear regression result also shows in the plot and the average residual was 0.9942. This plot confirms that the aperture intensity is linear to the actual light transmittance rate.

4.6 Distortion

Another problem that has been caused by the use of simple lenses in the relay optics is image distortion. The geometric distortion is calibrated by using the Matlab camera calibration toolbox as shown in **Fig. 6**. The circle indicates a center of distortion and the arrows represent displacements of the pixel introduced by the lens distortion. These calibrated camera parameters will be used to compensate the geometric distortions in the captured images.

4.7 PSF Evaluation

Lens aberration and diffraction may distort the actual PSFs. To assess the PSF quality of the prototype camera, we display a coded aperture and then calibrate the camera PSFs at 5 depths

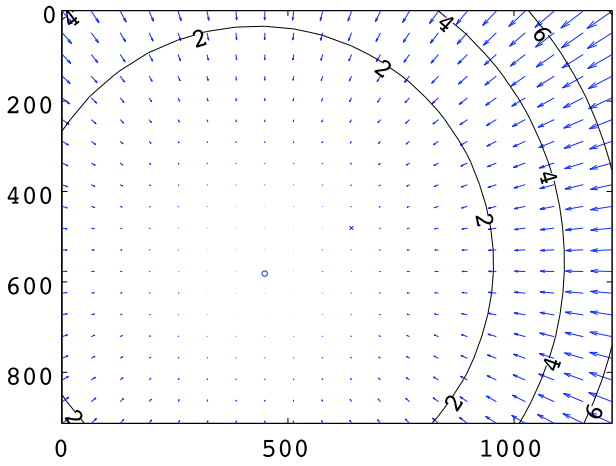


Fig. 6 Geometric distortion due to the use of simple optics.

and 5 different view angles. Without specific intentions, we use the aperture pattern as shown in **Fig. 7** (a) in this evaluation. Figure 7 (b) shows how PSFs varies with depth and field angle (image location in viewing angle). The PSFs were captured by the prototype camera with the aperture pattern as shown in Fig. 7 (a). We set a point light source on various positions in a scene so that we obtain the designated field angles and depths. The captured PSFs were normalized after cropping from an image, then were arranged them into Fig. 7 (b). We can see that the scale of PSF is related to the field angle. This is because the use of simple lenses in the relay optics leads to a field curvature.

We can see that the shapes of most PSFs are still very similar. To measure the similarity between these PSFs and the input aperture pattern, we normalize the scale of each PSF and compute its L_2 distance to the input pattern. A distance map is shown in the top of Fig. 7 (c). We can see that according to the L_2 distance, the PSF shape deviation decreases as the blur size increases. It is known that L_2 distance is not a good metric to measure the PSF similarities in defocus deblurring. To measure the dissimilarity between two PSFs, we use the Wiener reconstruction error when an image is blurred with one PSF and then deconvolved with another PSF. This reconstruction error turns out to be a variant of K-L divergence as shown in Ref. [22]. We plot this dissimilarity map in the bottom of Fig. 7 (c). We can see that all the dissimilarity values are small and decrease as the blur size decrease.

The specifications of the prototype programmable aperture camera are shown in **Table 2** as a summary.

5. Evaluation by Applications

5.1 Programmable Aperture for Defocus Deblurring

Another important limit of most existing coded aperture implementations is that the actual shape of the produced PSF often deviates from the input pattern due to lens aberration and diffraction. Note that the effects of lens aberration and diffraction can be quite different in different lenses. For the complexity of the modern lenses, it is difficult to take these effects into account during pattern optimization. The effects of these imperfections on the optimality of the apertures are often overlooked in the literature.

With a programmable aperture camera, we will be able to evaluate the input aperture pattern by analyzing the captured images,

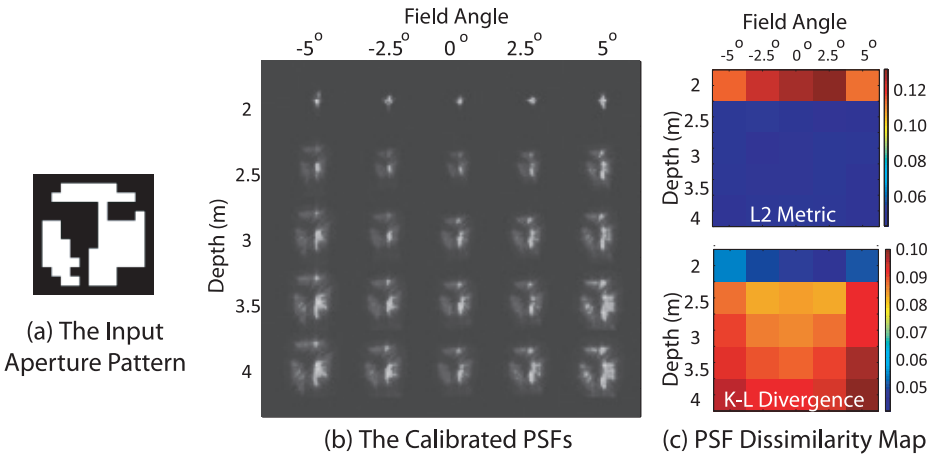


Fig. 7 Evaluating the PSFs of the prototype camera. (a) The coded aperture pattern used in the evaluation. This pattern is picked without specific intentions. (b) The calibrated PSFs at five depths ranging from 2 m to 4 m, and five field angles (image location in viewing angle) ranging from −5° to 5°. We can see that the scale of PSFs varies with both depth and field angle (due to field curvature), while the shape of PSFs appears similar. (c) The shape dissimilarity between the input pattern and each PSF is computed according to two metrics: L_2 distance at the top, and K-L divergence at the bottom (as used in Ref. [22]).

Table 2 Specification of the prototype camera.

Image resolution	1384 × 1036 pixels
Frame rate	25 fps
Minimum F-number	2.96
FOV(diagonal)	24.8° (25 mm Nikon C-mount)
Light transmittance	16.5%

and then improve the aperture patterns dynamically for a better performance. In this experiment, we apply this idea to the coded aperture technique for defocus deblurring.

Zhou and Nayar [2] propose a comprehensive criterion of aperture evaluation for defocus deblurring, which takes image noise level, the prior structure of natural images, and deblurring algorithm into account. They have also shown that the optimality of an aperture pattern can be different at different noise levels and scene settings. For a PSF k, its score at a noise level σ is measured as:

$$R(K|\sigma) = \Sigma \frac{\sigma^2}{|K|^2 + \sigma^2/|F_0|^2}, \tag{5}$$

where K is the Fourier transform of the PSF k, and F_0 is the Fourier transform of the ground truth focused image. This definition can be re-arranged as

$$\begin{aligned} R(K|\sigma) &= \Sigma \frac{\sigma^2 \cdot |F_0|^2}{|K|^2 \cdot |F_0|^2 + \sigma^2} \approx \Sigma \frac{\sigma^2 \cdot A}{|F|^2 + \sigma^2} \\ &\propto \Sigma \frac{A}{|F|^2 + \sigma^2}, \end{aligned} \tag{6}$$

where A is the average power spectrum of natural images as defined in Ref. [2], and F is the Fourier transform of the captured image. Therefore, given a captured defocused image F, Eq. (6) can be used to directly predict the quality of deblurring without calibrating the PSF and actually performing deblurring, while all the effects of aberrations and diffraction have been taken into account. Obviously, for the best deblurring quality, we should choose the aperture pattern which yields the lowest R value. The detailed discussion of the aperture selection and theoretical background are in Zhou's paper [2].

In our experiment, we capture a set of defocused images of an IEEE resolution chart (shown in the first row of **Fig. 8**) by using the aperture patterns shown in Fig. 1. We compute the R value from each captured image and find that the lowest R value is achieved by using the pattern shown in Fig. 8 (e). This indicates that this pattern is the best among all these candidate patterns in the present imaging condition and scene settings.

Note that this prediction is made directly from the observed defocused images without PSF calibration or deblurring. The computation only involves few basic arithmetic operations and one Fourier transform, and therefore can be done at real time. For comparison, the second row of Fig. 8 shows the deblurring results of several different aperture patterns. These results confirm that the pattern in (e) is the best for defocus deblurring in this particular image condition.

5.2 Programmable Aperture for Depth from Defocus

We tested the coded aperture technique for depth from defocus (DFD) proposed by Zhou et al. [4]. It is well know that PSF, which has a lot of zero-crossings in a frequency domain, is preferable for good depth discrimination. Conversely, defocus deblurring requires broadband PSF. Hence, they proposed to use a coded aperture pair for having contradictive properties; one of the pair has zero-crossings, but both of them realize broadband. The programmable aperture camera can be easy to get images with the two different coded apertures.

If we assumed that f_0 is a latent all in focused image, images f_i captured by coded apertures k_i are modeled by:

$$f_i = f_0 \otimes k_i^{d^*} + \eta_i, \ i = 1, 2, \tag{7}$$

where d^* is an actual scene depth and $k_i^{d^*}$ indicate a scaled version of blurred kernel k_i corresponding to the depth d^*. η is an additive image noise. Equation (7) can be expressed in frequency domain as:

$$F_i = F_0 \cdot K_i^{d^*} + \zeta_i, \ i = 1, 2, \tag{8}$$

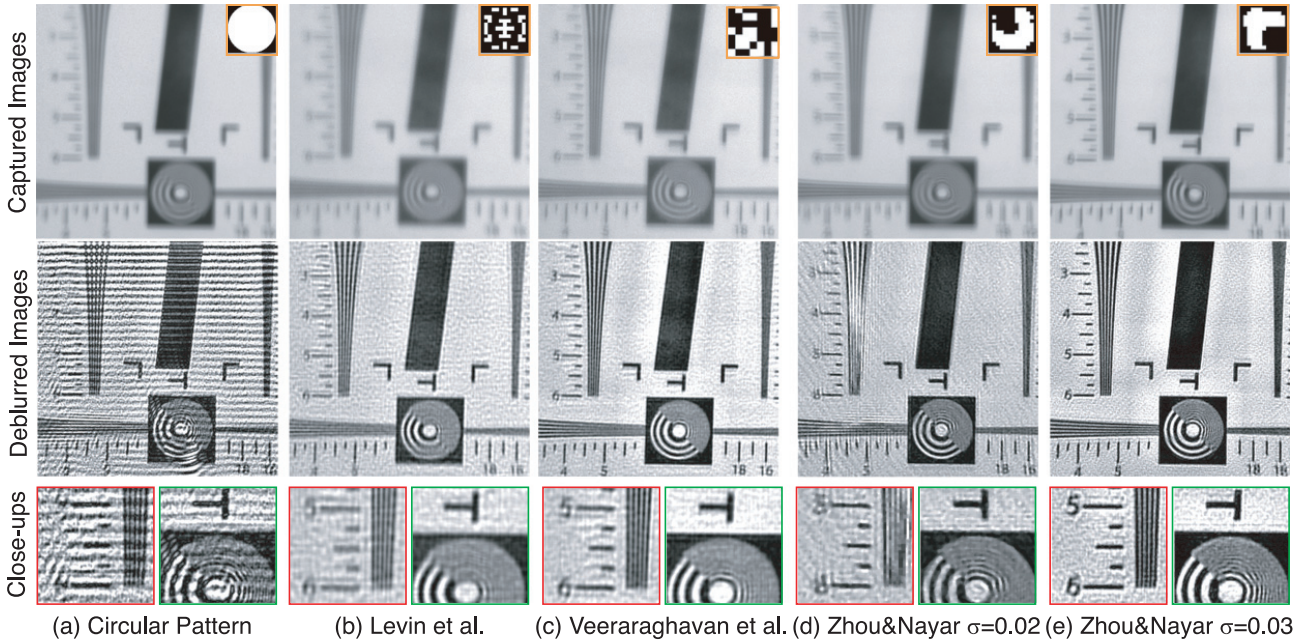


Fig. 8 Pattern selection for defocus deblurring by using the programmable aperture camera. We capture a set of defocused images of an IEEE resolution chart using the patterns shown in Fig. 1, and evaluate their qualities using Eq. (6). The pattern shown in Column (e) is found to be the best according to our proposed criterion. To verify this prediction, we calibrate the PSFs in all the captured images, do deblurring, and show deblurring results (the second and third rows). We can see that the deblurring result in Column (e) is the best, which is consistent with the prediction.

where F_i, F_0, $K_i^{d^*}$ and ζ are Fourier transform of f_i, f_0, $k_i^{d^*}$ and η, respectively. Our objective is to find the depth $\hat{d}$ and deblurred image $\hat{F}_0$. We define a objective function for solving this problem as:

$$W^{(d)} = \sum_{i=1,2} |IFFT(\hat{F_0}^{(d)} * K_i^{\hat{d}} - F_i)|, \tag{9}$$

$$\hat{F_0} = \frac{F_1 \cdot \bar{K_1}^{\hat{d}} + F_2 \cdot \bar{K_2}^{\hat{d}}}{|K_1^{\hat{d}}|^2 + |K_2^{\hat{d}}|^2 + |C|^2},$$

where $IFFT$ is the 2D inverse Fourier transform. $\bar{K}$ is the complex conjugate of K and $|K|^2 = K \cdot \bar{K}$. C is a reciprocal of signal to noise ratio of the captured images. You can find the detailed descriptions or derivations is in Ref. [4]. By minimizing $W^{(d)}(x, y)$ for each pixel, we can obtain the depth map U as:

$$U(x, y) = \arg\min_{d \in D} W^{(d)}(x, y). \tag{10}$$

Also we can find the all in focus image I from the estimated depth map as:

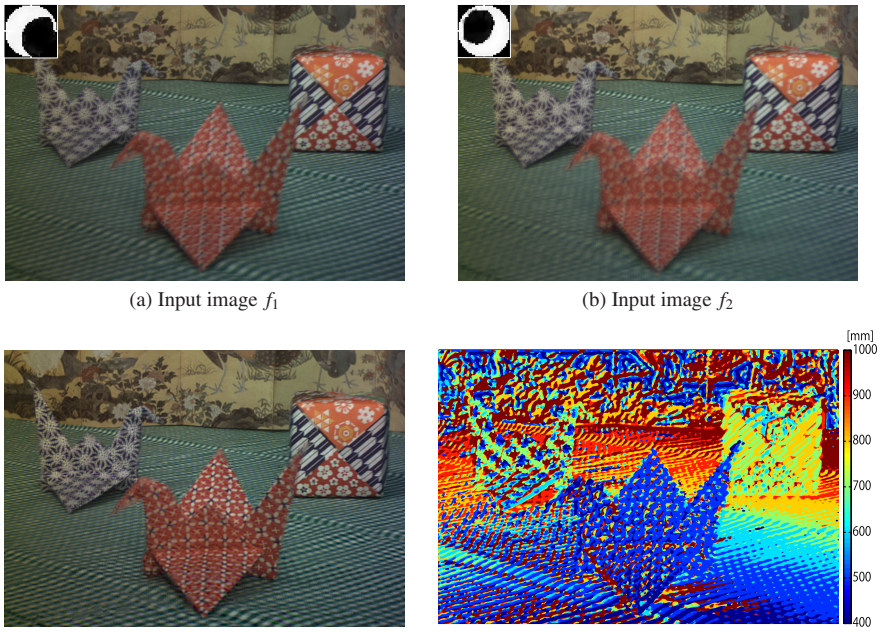
$$I(x, y) = \hat{F_0}^{(U_{x,y})}(x, y). \tag{11}$$

We carried out an experiment of DFD application by using this algorithm. We captured a scene which has three object and backdrop with different depths (a red crane was on 500 mm, a blue crane and cube were on 700 mm and a backdrop was on 1,000 mm from the camera) as shown in **Fig. 9**. Figure 9 (a), (b) show the input images captured by a coded aperture pair. The insets on top left of the images indicate the shapes of coded apertures when the images were captured. Figure 9 (c), (d) show the estimated results of the deblurred image and the depth map. **Figure 10** shows a similar experimental result by using conventional circular apertures with different radii for comparison. Figure 10 (d) shows the circular DFD cannot estimate a scene depth. Notice that we did not use any post-processing, such as BP or Graph Cut etc. and the both depth maps were raw pixel-wise estimations described by Eq. (10), so that we can easy to understand the coding difference. As a result, there are a lot of artifacts, ringing edge and remaining blur, in the deblurred image as shown in Fig. 10 (c), since the image were deblurred by different size of kernels. On the other hand, you can see that in-focused radiance of the scene and the scene depths were correctly recovered as shown in Fig. 9 (c), (d). We confirmed that prototype camera can be applicable for DFD application and the coded aperture pair has an advantage to the conventional one. In Ref. [4], they replaced the SLR lenses with different aperture masks for taking input images of DFD. The programmable aperture camera is feasible for the DFD application, since the camera easy and quickly changes the patterns at video rate.

5.3 Programmable Aperture for Light Field Acquisition

We finally use our prototype programmable aperture camera to re-implement the multiplexing light field acquisition method, which is first proposed by Liang et al. [5] A 4D light field is often represented as $l(u, v, x, y)$ [23], where (u, v) is the coordinates on the aperture plane and (x, y) is the coordinates in the image plane.

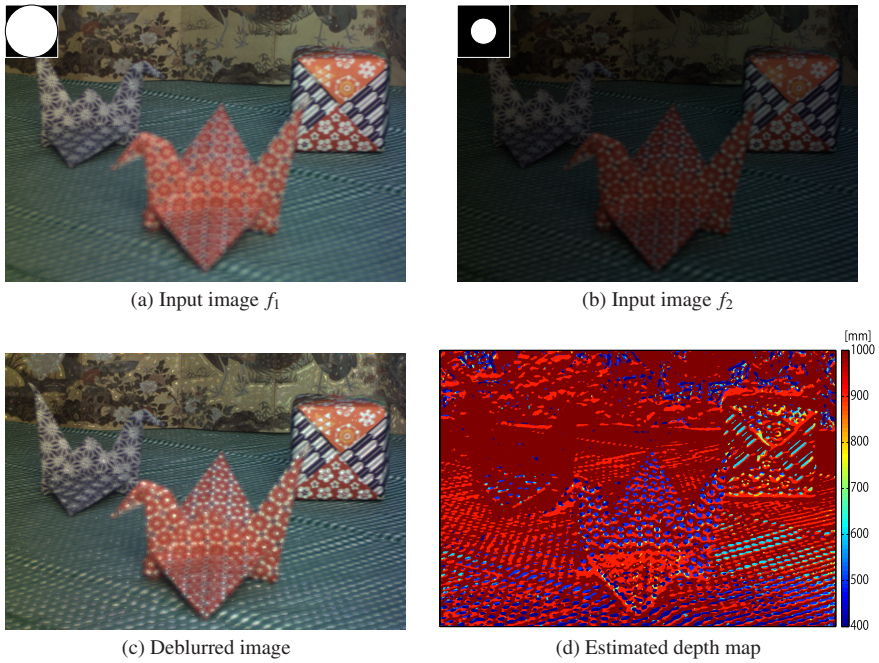
For a light field acquisition technique using coded aperture, the spatial resolution in the (x, y) space is simply determined by the sensor resolution and the angular resolution in the (u, v) space is determined the resolution of coded apertures. Bando et al. [8] use a 2×2 color coded aperture to capture light fields and then use the information to do layer estimation and matting. Liang et al. [5] propose a multiplexing technique to capture light fields up to 7×7 angular resolution. For any $m \times n$ angular resolution light



(a) Input image f_1 (b) Input image f_2

(c) Deblurred image (d) Estimated depth map

Fig. 9 Depth from defocus by using coded aperture pair. Input images (a), (b) captured by coded apertures indicated on the insets. The deblurred image (c) and depth map (d) are estimated by the proposed method.



(a) Input image f_1 (b) Input image f_2

(c) Deblurred image (d) Estimated depth map

Fig. 10 DFD by a conventional circular aperture pair.

field acquisition, the multiplexing method requires $m \times n$ images captured using $m \times n$ different coded apertures.

With our prototype programmable aperture camera, it is easy to capture light fields with various angular resolutions. We use S-matrix for the multiplexing coding (see Ref. [24] for a deep discussion on the multiplexing coding). **Figure 11** (top) shows four of the 31 aperture patterns [*5] that we generate from an S-Matrix. Since the aperture pattern of the prototype camera can be updated at a video frame rate (25 fps), it only takes 1.2 seconds to capture all of the images. If we could increase the camera frame rate further or lower the aperture resolution, the programmable aperture camera could be able to capture light fields of moving objects.

From the 31 captured images, we recover the light field of resolution $1280 \times 960 \times 31$ (7×5 (u, v) resolution excluding the four corners). **Figure 12** shows the images for different viewpoints (u, v) and their close-ups. From the close-ups, we can see the disparities of the text clearly. With the recovered light field, people will be able to do further post-processing including depth estimation and refocusing as shown in Refs. [5], [8].

[*5] This is because the code length of S-matrix must be $2^n - 1$.

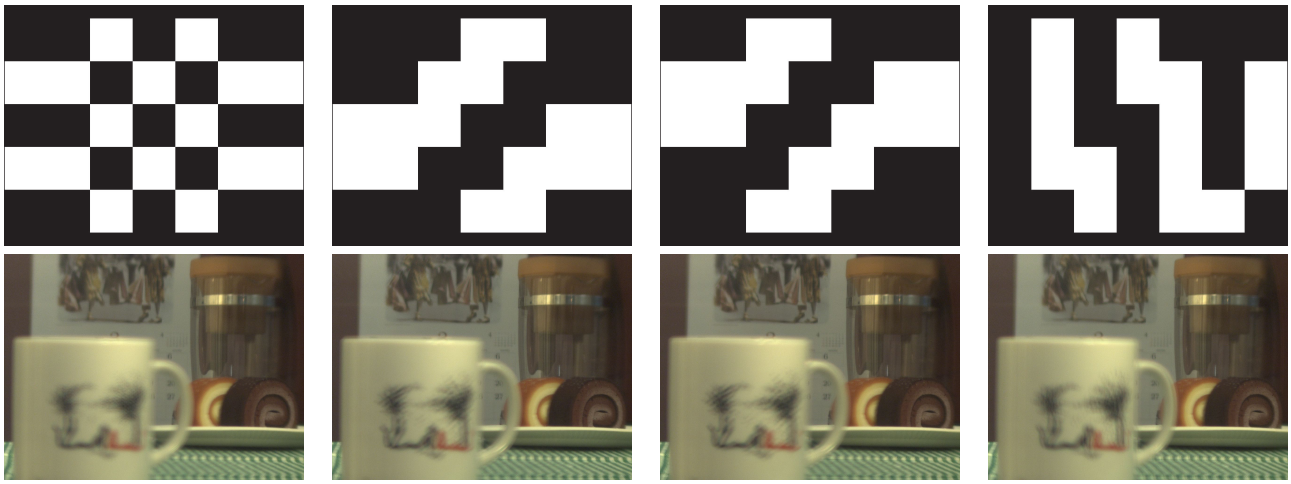
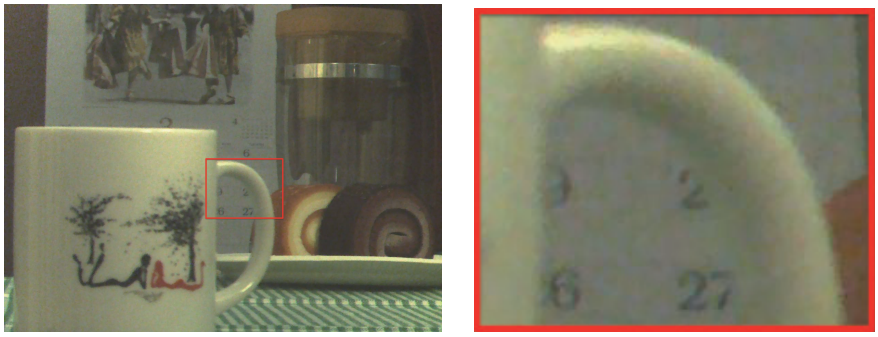
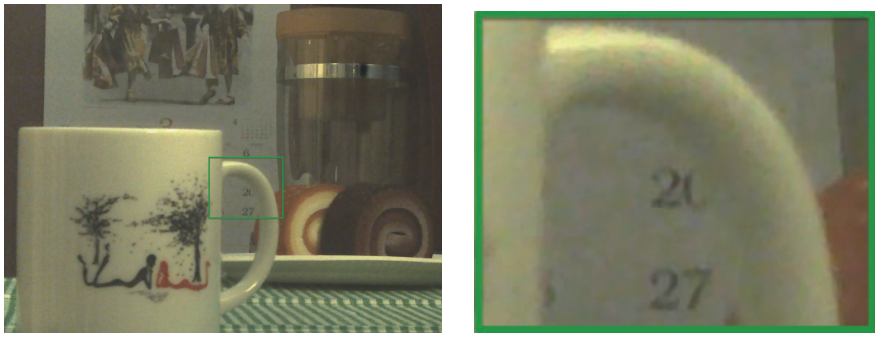


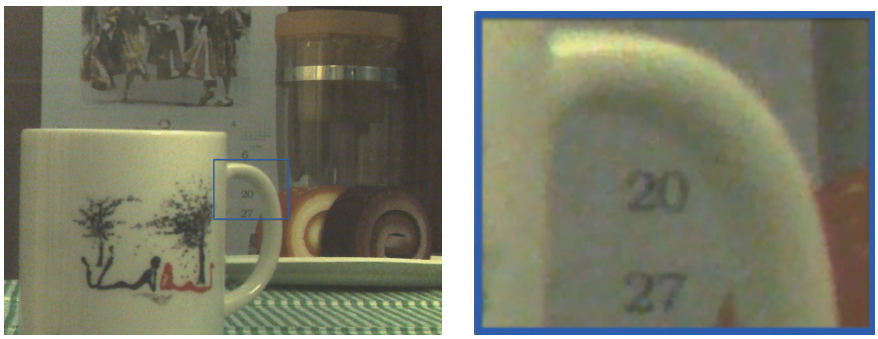
Fig. 11 Four multiplexing aperture codings and the corresponding captured images. Upper row shows four of the 31 aperture patterns that we generate from an S-Matrix. Bottom row shows the four corresponding captured images.



$(u, v) = (2, 3)$ Close-up



$(u, v) = (4, 3)$ Close-up



$(u, v) = (6, 3)$ Close-up

Fig. 12 The reconstructed 4D light field. Images from three different view points (u, v) are generated from the reconstructed 4D light field, and their close-ups are shown in thier right. From the close-up images, we can see the disparities that the numbers, 20 and 27, on the backside calender is moving left w.r.t. the frontal mag handle position.

6. Conclusion and Perspectives

In this paper, we propose to build a programmable aperture camera using an LCoS device which enables us to implement aperture patterns of high brightness contrast, light efficient and resolution at a video frame rate. Another important feature of this design is that any C-Mount or F-Mount lenses can be easily attached to the proposed camera without being disassembled. These features make our design applicable to a variety of coded aperture techniques. We demonstrate the use of our proposed programmable aperture camera in three applications: defocus deblurring, coded aperture pair for DFD, and multiplexing light field acquisition.

Acknowledgments This work was supported by KAKENHI 21680018 (Grant-in-Aid for Young Scientists (A)) and KAKENHI 20650023 (Grant-in-Aid for challenging Exploratory Research).

Reference

[1] Veeraraghavan, A., Raskar, R., Agrawal, A., Mohan, A. and Tumblin, J.: Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing, *ACM Trans. Graphics*, Vol.26, No.3, Article 69 (2007).

[2] Zhou, C. and Nayar, S.: What are good apertures for defocus deblurring?, *International Conference of Computational Photography*, San Francisco, U.S. (Apr. 2009).

[3] Levin, A., Fergus, R., Durand, F. and Freeman, W.: Image and depth from a conventional camera with a coded aperture, *ACM Trans. Graphics*, Vol.26, No.3, Article 70 (2007).

[4] Zhou, C., Lin, S. and Nayar, S.: Coded Aperture Pairs for Depth from Defocus, *Proc. International Conference on Computer Vision*, Kyoto, Japan, pp.325–332 (2009).

[5] Liang, C.K., Lin, T.H., Wong, B.Y., Liu, C. and Chen, H.: Programmable aperture photography: Multiplexed light field acquisition, *ACM Trans. Graphics*, Vol.27, No.3, Article 55 (2008).

[6] Raskar, R., Agrawal, A. and Tumblin, J.: Coded exposure photography: Motion deblurring using fluttered shutter, *ACM Trans. Graphics*, Vol.25, No.3, pp.795–804 (2006).

[7] Gottesman, S. and Fenimore, E.: New family of binary arrays for coded aperture imaging, *Applied Optics*, Vol.28, No.20, pp.4344–4352 (1989).

[8] Bando, Y., Chen, B. and Nishita, T.: Extracting depth and matte using a color-filtered aperture, *ACM Trans. Graphics*, Vol.27, No.5, Article 134 (2008).

[9] Hasinoff, S., Kutulakos, K., Durand, F. and Freeman, W.: Time-constrained photography, *Proc. International Conference on Computer Vision*, pp.333–340 (2009).

[10] Caroli, E., Stephen, J., Cocco, G., Natalucci, L. and Spizzichino, A.: Coded aperture imaging in X-and gamma-ray astronomy, *Space Science Reviews*, Vol.45, No.3-4, pp.349–403 (1987).

[11] Welford, W.: Use of annular apertures to increase focal depth, *Journal of the Optical Society of America A*, Vol.50, No.8, pp.749–753 (1960).

[12] Mino, M. and Okano, Y.: Improvement in the OTF of a defocused optical system through the use of shaded apertures, *Applied Optics*, Vol.10, No.10, pp.2219–2225 (1971).

[13] Varamit, C. and Indebetouw, G.: Imaging properties of defocused partitioned pupils, *Journal of the Optical Society of America A*, Vol.2, No.6, pp.799–802 (1985).

[14] Ojeda-Castafieda, J., Andres, P. and Diaz, A.: Annular apodizers for low sensitivity to defocus and to spherical aberration, *Optics Letters*, Vol.11, No.8, pp.487–489 (1986).

[15] Zomet, A. and Nayar, S.: Lensless imaging with a controllable aperture, *Proc. Computer Vision and Pattern Recognition*, pp.339–346 (2006).

[16] Aggarwal, M. and Ahuja, N.: Split Aperture Imaging for High Dynamic Range, *International Journal of Computer Vision*, Vol.2, pp.10–17 (2001).

[17] Green, P., Sun, W., Matusik, W. and Durand, F.: Multi-aperture photography, *ACM Trans. Graphics*, Vol.26, No.3, Article 68 (2007).

[18] Dowski, E.R. and Cathey, W.T.: Extended Depth of Field through Wave-front coding, *Applied Optics*, Vol.34, No.11, pp.1859–1866 (1995).

[19] Wikipedia: Liquid crystal on silicon. available from ⟨http://en.wikipedia.org/wiki/Liquid_crystal_on_silicon⟩.

[20] Mannami, H., Sagawa, R., Mukaigawa, Y., Echigo, T. and Yagi, Y.: High dynamic range camera using reflective liquid crystal, *Proc. International Conference on Computer Vision*, pp.1–8 (2007).

[21] Nayar, S.K., Branzoi, V. and Boult, T.: Programmable imaging: Towards a flexible camera, *International Journal of Computer Vision*, Vol.70, pp.7–22 (2006).

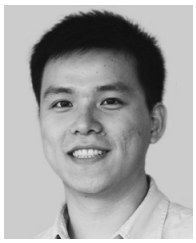
[22] Nagahara, H., Kuthirummal, S., Zhou, C. and Nayar, S.: Flexible depth of field photography, *Proc. European Conference on Computer Vision*, Vol.4, No.LNCS5305, pp.60–73 (2008).

[23] Levoy, M. and Hanrahan, P.: Light field rendering, *Proc. ACM SIGGRAPH*, pp.31–42 (1996).

[24] Schechner, Y., Nayar, S. and Belhumeur, P.: A theory of multiplexed illumination, *Proc. International Conference on Computer Vision*, Vol.2, pp.808–815 (2003).

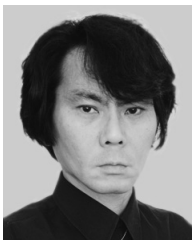
[25] Nagahara, H., Zhou, C., Watanabe, T., Ishiguro, H. and Nayar, S.: Programmable Aperture Camera Using LCoS, *Proc. European Conference on Computer Vision*, Vol.2, No.LNCS6316, pp.337–350 (2010).

Hajime Nagahara received his B.E. and M.E. degrees in Electrical and Electronic Engineering from Yamaguchi University in 1996 and 1998, respectively, and Ph.D. degree in system engineering from Osaka University in 2001. He was a research associate of the Japan Society for the Promotion of Science (2001–2003). He was Assistant Professor at the Graduate School of Engineering Science, Osaka University, Japan (2003–2010). He was Visiting Associate Professor at CREA University of Picardie Jules Verns, France, in 2005. He was a visiting researcher at Columbia University in 2007–2008. Since 2010, he has been Associate Professor in Faculty of Information Science and Electrical Engineering at Kyushu University. Computational photography, computer vision, and virtual reality are his research areas. He received the Honorable Mention Award of ACM VRST in 2003.



Changyin Zhou received his B.S. degree in Statistics and M.S. 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 IEEE.

Takuya Watanabe recieved his B.E. and M.E. degrees in Enginerring Science from Osaka University in 2009 and 2011, respectively. He has studied in a programmable aperture camera project while he was in the school. He is currently in Denso Co.



Hiroshi Ishiguro received his D.Eng. in Systems Engineering from the Osaka University, Japan in 1991. He is currently Professor of Department of Systems Innovation in the Graduate School of Engineering Science at Osaka University (2009–). He is also Group Leader (2011–) of Hiroshi Ishiguro Laboratory at the Advanced Telecommunications Research Institute, where he is ATR Fellow (2010–). He was previously Research Associate (1992–1994) in the Graduate School of Engineering Science at Osaka University and Associate Professor (1998–2000) in the Department of Social Informatics at Kyoto University. He was also Visiting Scholar (1998–1999) at the University of California, San Diego, USA. He was Associate Professor (2000–2001) and Professor (2001–2002) in the Department of Computer and Communication Sciences at Wakayama University. He then moved to the Department of Adaptive Machine Systems in the Graduate School of Engineering Science at Osaka University as a Professor (2002–2009). His research interests include distributed sensor systems, interactive robotics, and android science.

Shree K. Nayar received his Ph.D. 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, CVPR 2004 and ICCP 2010. 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 (in 2006), and the Carnegie Mellon Alumni Achievement Award (in 2009). He was elected to the National Academy of Engineering in 2008 and to the American Academy of Arts and Sciences in 2011.

(Communicated by *Guy Godin*)