

# Cauchy Aperture and Perfect Reconstruction Filters for Extending Depth-of-Field from Focal Stack

Akira KUBOTA<sup>†a)</sup>, Kazuya KODAMA<sup>††</sup>, Members, and Asami ITO<sup>†</sup>, Nonmember

**SUMMARY** A pupil function of aperture in image capturing systems is theoretically derived such that one can perfectly reconstruct all-in-focus image through linear filtering of the focal stack. The perfect reconstruction filters are also designed based on the derived pupil function. The designed filters are space-invariant; hence the presented method does not require region segmentation. Simulation results using synthetic scenes shows effectiveness of the derived pupil function and the filters.

**key words:** extended depth-of-field, focal stack, aperture, perfect reconstruction filters, image fusion

## 1. Introduction

Depth-of-Field (DOF) of an imaging system is the depth range in which scenes appear in-focus in the captured image. Most imaging systems, especially in microscopy, have narrow DOF; hence the captured images suffer from blurring in the regions which depths are not in the DOF. In order to obtain the all-in-focus images where all the regions are in-focus, extensive methods for extending DOF have been studied in the last decades. These methods are categorized into two groups: image recovery from a single captured image and image fusion using multiple captured images.

In image recovery approaches, so called blind image recovery [1], an all-in-focus image is to be recovered from a single captured image. The captured image, say  $g(x, y)$ , is modeled as

$$g(x, y) = \iint h(x, y, p, q) f(p, q) dpdq, \quad (1)$$

where  $f$  denotes the desired all-in-focus image,  $h$  point-spread-function (PSF) and  $(x, y)$  image coordinates. Since PSF  $h$  is unknown and space-variant (varies with scene depth), it is generally difficult to estimate  $f$  with adequate quality for arbitrary scenes without some priors on image  $f$ .

Recently, to robustly recover  $f$ , some techniques [2]–[5] have been presented based on designs of optics in imaging system. These methods can make PSFs space-invariant by putting special optical elements at the aperture plane. The model of the imaging process is expressed by convolution of  $f$  and  $h$ :

Manuscript received February 17, 2019.

Manuscript revised June 9, 2019.

Manuscript publicized August 16, 2019.

<sup>†</sup>The authors are with Chuo University, Tokyo, 112–8551 Japan.

<sup>††</sup>The author is with National Institute of Informatics, Tokyo, 101–8430 Japan.

a) E-mail: kubota@elect.chuo-u.ac.jp

DOI: 10.1587/transinf.2019PCP0006

$$g(x, y) = \iint h(x - p, y - q) f(p, q) dpdq; \quad (2)$$

hence the all-in-focus image  $f$  can be recovered simply by deconvolving  $g$  without involving depth estimation.

This space-invariant imaging can be also achieved by focus sweep methods [6]–[8], where an image was captured over moving the sensor plane during the exposure time. The PSF of the captured image, called integrated PSF, is shown to be nearly space-invariant and therefore the imaging model is expressed by Eq. (2).

The other category is image fusion [9]–[16]. Unlike above deconvolution methods using a single captured image, image fusion methods capture multiple images with different focus depths (a set of them is called focal stack) and fuse them into the all-in-focus image. Because scene textures at any depth appear in-focus in one of the focal stack images, much robust recovery is possible. Image fusion methods consist of two parts. The first part is to select dominant values from pixel values in the spatial domain or dominant coefficients in the transformed domain. The second part is to combine these selected values/coefficients into the fused image. These methods are essentially same as region segmentation; hence often suffer from blocking artifacts due to incorrect regions estimations. In addition, a key to high-quality fusion is the selection rule [17], based on which true values/coefficients of the all-in-focus image can be correctly selected; however, it is difficult to design the selection rule for perfect reconstruction.

To tackle these problems above, this paper presents a novel filter-based image fusion method for reconstructing an all-in-focus image  $f$  from focal stack, say  $\{g_1, g_2, \dots, g_N\}$ , by

$$f(x, y) = \sum_{i=1}^N k_i(x, y) * g_i(x, y), \quad (3)$$

where the operation  $*$  denotes two-dimensional convolution and the set of filters  $k_i$  ( $i = 1, 2, \dots, N$ ) is the synthetic filter bank. In this paper, the pupil function of the aperture in image capturing systems is derived such that one can perfectly reconstruct all-in-focus image by Eq. (3) in theory. The filter bank is also designed based on the derived pupil function. The designed filters are all space-invariant; hence the presented method requires neither depth estimation nor region segmentation.

For the case of  $N = 2$ , the author derived synthetic filters that perfectly reconstruct all-in-focus image [18]. It

is possible to show the existence of the filters for normally used pupil functions such as Gaussian and pill-box functions. Extending this idea to the case of  $N \geq 3$ , however, does not lead stable filters; hence an iterative reconstruction method [19], [20] and regularization method [21] have been applied for this case. But, they estimate focus slice images (in-focus regions) and do not derive filters that perfectly reconstruct the all-in-focus image directly from the focal stack. In this paper, to achieve perfect reconstruction, the pupil function is theoretically designed. In practice, the focal stack should be captured by a specially designed camera with the aperture of the presented pupil function.

Some filtering approaches to all-in-focus image reconstruction have been presented. The authors [22], [23] presented three-dimensional deconvolution method in the Frequency domain that recovers the all-in-focus image from the focal stack. They [25] and Levin et al. [26] showed that this approach is achieved by two-dimensional filtering of the averaged image of the focal stack, which is the same with focus sweep method. However, the imaging models of the focal stack or its average used in these methods are approximations of the accurate version; for robust reconstruction, the focusing range must be set wider than the actual depth range of the scene.

Designing pupil functions of the aperture, called coded aperture, has been recently studied in depth-from-defocus methods [27]–[29] that use a few images captured with different focus depths to reconstruct shape of scenes as well as all-in-focus image. But, in image fusion methods, effect of pupil function and its optimal design have not been well studied.

This paper is an extended version of the paper in [30]. In this paper, the performance evaluations of reconstruction accuracy and noise sensitivity are additionally discussed in detail.

## 2. Cauchy Aperture and Perfect Reconstruction Filters

### 2.1 Imaging Model

#### 2.1.1 Focal Stack Model

A focal stack is a set of images captured by changing the distance of the imaging plane from the lens with equal interval. Let  $g_i(x, y)$  be the image captured when the distance is  $v_i$  ( $i = 1, 2, \dots, N$ ). The coordinate  $(x, y)$  represents the image coordinate perpendicular to the optical axis. Here we assume that  $v_1 > v_2 > \dots > v_N$  holds and the magnification difference is already corrected among the captured images.

$$H = \begin{pmatrix} 1 & A(t\xi, t\eta; \sigma) & A(2t\xi, 2t\eta; \sigma) & \dots & A((N-1)t\xi, (N-1)t\eta; \sigma) \\ A(t\xi, t\eta; \sigma) & 1 & A(t\xi, t\eta; \sigma) & \dots & A((N-2)t\xi, (N-2)t\eta; \sigma) \\ A(2t\xi, 2t\eta; \sigma) & A(t\xi, t\eta; \sigma) & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & A(t\xi, t\eta; \sigma) \\ A((N-1)t\xi, (N-1)t\eta; \sigma) & A((N-2)t\xi, (N-2)t\eta; \sigma) & \dots & A(t\xi, t\eta; \sigma) & 1 \end{pmatrix}.$$

The pupil function is assumed to be circularly symmetric and expressed by  $a(x, y; \sigma)$  with a scaling parameter  $\sigma$  that determines the amount of blur degree. It should satisfy the condition of

$$\begin{cases} \iint a(x, y; \sigma) dx dy = 1 \\ a(x, y; \sigma) \geq 0 \end{cases}. \quad (4)$$

Let  $f_i(x, y)$  be the in-focus regions in the captured image  $g_i(x, y)$  and call it focus slice image. The focal stack is modeled by combination of the focus slices [18]–[21], [31] as

$$g_i(x, y) = \sum_{j=1}^N h_{ij}(x, y) * f_j(x, y), \quad i = 1, 2, \dots, N, \quad (5)$$

where  $*$  denotes two-dimensional convolution and the function  $h_{ij}(x, y)$  represents a point spread function (PSF), which is a scaled version of the pupil function

$$h_{ij}(x, y) = \begin{cases} \frac{1}{|i-j|t^2} a\left(\frac{x}{|i-j|t}, \frac{y}{|i-j|t}; \sigma\right) & (i \neq j) \\ \delta(x, y) & (i = j) \end{cases}. \quad (6)$$

The parameter  $t$  is a ratio of the distance interval of the image plane to the focus length and  $\delta(x, y)$  is the two-dimensional Dirac delta function.

The model in Eq. (5) is represented in the Fourier domain in matrix-vector form:

$$\mathbf{g} = H\mathbf{f}, \quad (7)$$

where

$$\mathbf{g} = \begin{pmatrix} G_1(\xi, \eta) \\ G_2(\xi, \eta) \\ \vdots \\ G_N(\xi, \eta) \end{pmatrix}, \quad \mathbf{f} = \begin{pmatrix} F_1(\xi, \eta) \\ F_2(\xi, \eta) \\ \vdots \\ F_N(\xi, \eta) \end{pmatrix},$$

and  $H$  (see equation at the bottom).

In the above equation, functions  $G_i$ ,  $F_i$  and  $A$  denote the Fourier transform of  $g_i$ ,  $f_i$  and  $a$ , respectively, and  $(\xi, \eta)$  is spatial frequencies. Note that the matrix  $H$  is a symmetric Toeplitz matrix.

#### 2.1.2 All-in-Focus Image Model

The all-in-focus image  $f(x, y)$  is represented by a sum of the focus slice images  $f_i(x, y)$ :

$$f(x, y) = \sum_{i=1}^N f_i(x, y). \tag{8}$$

Its Fourier transform is written in vector form by

$$F(\xi, \eta) = \mathbf{1}_N^T \mathbf{f}, \tag{9}$$

where  $F$  is the Fourier transform of  $f$ ,  $\mathbf{1}_N$  is defined by a  $N$ -dimensional vector which elements are all 1, and  $T$  denotes the transpose operation.

### 2.2 Deriving Synthetic Filters and Pupil Function

Eliminating the focus slice images  $\mathbf{f}$  from Eqs. (7) and (9) yields

$$F(\xi, \eta) = (\mathbf{1}_N^T H^{-1}) \mathbf{g}. \tag{10}$$

If the column vector  $\mathbf{1}_N^T H^{-1}$  exists, each element gives the frequency characteristic  $K_i$  of the filter  $k_i$  to the image  $g_i$ .

A sufficient condition such that  $\mathbf{1}_N^T H^{-1}$  exists is that the matrix  $H$  should become the form of Kac-Murdock-Szegö matrix [32] (see the matrix  $H_{KMS}$  at the bottom). In short, since  $H$  is symmetric, this condition is written by

$$A(|i-j|t\xi, |i-j|t\eta; \sigma) = A^{|i-j|}(t\xi, t\eta; \sigma). \tag{11}$$

If  $H$  satisfies this condition,  $H^{-1}$  exists for the case of  $A(t\xi, t\eta; \sigma) \neq 1$  (as shown in the matrix  $H^{-1}$  at the bottom) and Eq. (10) can be represented by

$$F(\xi, \eta) = \frac{1}{1 + A(t\xi, t\eta; \sigma)} (G_1 + G_N) + \frac{1 - A(t\xi, t\eta; \sigma)}{1 + A(t\xi, t\eta; \sigma)} (G_2 + \dots + G_{N-1}). \tag{12}$$

It is found that the frequency characteristics of the filters are obtained to be

$$K_i(\xi, \eta) = \begin{cases} \frac{1}{1 + A(t\xi, t\eta; \sigma)} & (i = 1, N) \\ \frac{1 - A(t\xi, t\eta; \sigma)}{1 + A(t\xi, t\eta; \sigma)} & (i = 2, 3, \dots, N-1) \end{cases} \tag{13}$$

$$H_{KMS} = \begin{pmatrix} 1 & A(t\xi, t\eta; \sigma) & A^2(t\xi, t\eta; \sigma) & \dots & A^{N-1}(t\xi, t\eta; \sigma) \\ A(t\xi, t\eta; \sigma) & 1 & A(t\xi, t\eta; \sigma) & \dots & A^{N-2}(t\xi, t\eta; \sigma) \\ A^2(t\xi, t\eta; \sigma) & A(t\xi, t\eta; \sigma) & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & A(t\xi, t\eta; \sigma) \\ A^{N-1}(t\xi, t\eta; \sigma) & \dots & \dots & A(t\xi, t\eta; \sigma) & 1 \end{pmatrix}$$

$$H^{-1} = \frac{1}{1 - A^2(t\xi, t\eta; \sigma)} \begin{pmatrix} 1 & -A(t\xi, t\eta; \sigma) & & & \\ -A(t\xi, t\eta; \sigma) & 1 + A^2(t\xi, t\eta; \sigma) & -A(t\xi, t\eta; \sigma) & & \\ & & \ddots & \ddots & \\ & & & \ddots & \\ -A(t\xi, t\eta; \sigma) & 1 + A^2(t\xi, t\eta; \sigma) & -A(t\xi, t\eta; \sigma) & & \\ & & -A(t\xi, t\eta; \sigma) & 1 & \end{pmatrix}$$

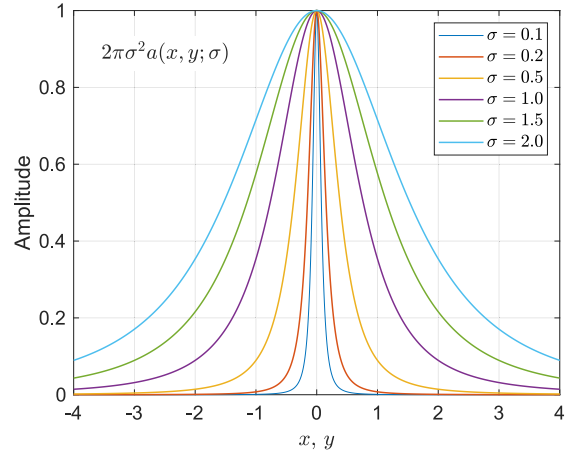


Fig. 1 Characteristics of Cauchy aperture function (the maximum amplitude is normalized to 1).

and that they are stable even for the case of  $A(t\xi, t\eta; \sigma) = 1$ . Therefore, all the filters exist for all the frequencies and are stable reconstruction filters.

The function  $A$  that satisfies the condition (11) in addition to the condition (4) is an exponential function of

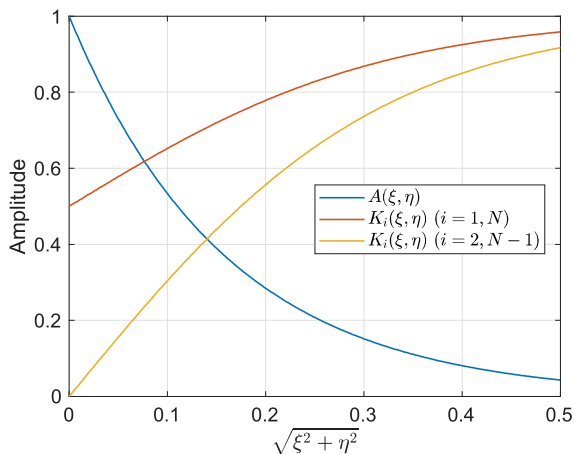
$$A(\xi, \eta; \sigma) = \exp \left\{ -2\pi\sigma \sqrt{\xi^2 + \eta^2} \right\}; \tag{14}$$

hence its inverse Fourier transform gives the pupil function in spatial domain as

$$a(x, y; \sigma) = \frac{\sigma}{2\pi(x^2 + y^2 + \sigma^2)^{3/2}}. \tag{15}$$

This is two-dimensional Cauchy distribution function [33]. We call this pupil function Cauchy aperture function in this paper. The examples of the function are plotted in Fig. 1.

Figure 2 shows the frequency characteristics of the filters for  $\sigma = 1.0$ . This indicates that all the filters work as high-pass filters and extract higher frequency components corresponding to the in-focus regions from the focal stack to reconstruct the all-in-focus image. The direct current and lower frequency components of the all-in-focus image were compensated from the first and the last focal stack images.



**Fig. 2** Frequency characteristics of synthetic filters and Cauchy PSF of  $\sigma = 1.0$ .

### 2.3 Advantages and Disadvantages

The advantage of the presented method is that the presented method can reconstruct all-in-focus image perfectly without requiring region segmentation, which means it is independent of the target scenes.

The presented method has mainly two disadvantages. The first is that noise in the focal stack images are accumulated in the reconstructed images. This effect will be theoretically and numerically evaluated in Sect. 3. The other is that the imaging model of focal stack images is in practice not correct for the occluding boundaries; therefore the presented method causes errors for those regions when using real captured focal stack images.

## 3. Simulation

### 3.1 Preparation of Ground Truth and Focal Stack Images

Three test images of “Baboon”, “Lenna” and “Peppers” were used as the ground truth of all-in-focus image  $f$ . These test images are 24-bit color images with  $512 \times 512$  pixels. The focal stack images  $g_i$  ( $i = 1, 2, \dots, N$ ) were synthetically generated based on the imaging model (7) using the focus slice images  $f_i$  which were created by equally dividing the ground truth into  $N$  thin rectangles from the left edge.

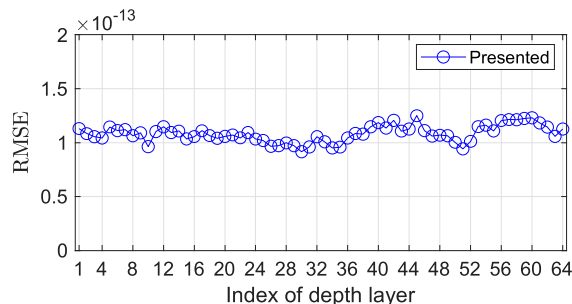
In this simulation, the maximum of the blur amount in the focal stack images,  $(N-1)\sigma t$ , was fixed to 10 [pixels], which means the depth range of the synthetic scene is fixed. Under this condition, the number of depth layers,  $N$ , were changed to 8, 16, 32 and 64. (See some examples of the focal stack images generated in Figs. 4 and 5)

### 3.2 Evaluation of Perfect Reconstruction

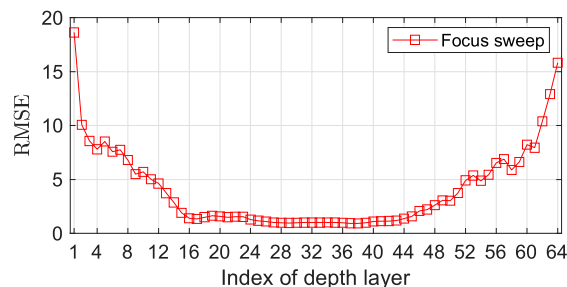
To precisely evaluate whether the presented method achieves perfect reconstruction, we calculated root mean squared error (RMSE) of the reconstructed images in double

**Table 1** RMSEs of reconstructed all-in-focus images in double precision for “Peppers” test image.

Method	Number of depth layers, $N$			
	8	16	32	64
Presented ( $\times 10^{-13}$ )	1.09	1.13	1.10	1.09
focus sweep [25], [26] ( $\times 1$ )	23.7	6.35	5.57	5.56



(a) Presented method



(b) Focus sweep method

**Fig. 3** RMSEs of focus slices for “Peppers” test image.

precision floating-point number. In this evaluation, we used the focal stack images that were stored in double precision without 8-bit quantization.

The results for “Peppers” image are shown in Table 1. All the RMSEs are quite small and less than  $10^{-12}$ , indicating that perfect reconstruction can be achieved by the presented method. For comparison, RMSEs of the focus sweep method were also calculated in double precision. As shown in Table 1, the RMSEs are much larger in comparison to those of the presented method. This shows that the focus sweep method cannot perfectly reconstruct all-in-focus images.

Figure 3 (a) and (b) show RMSEs of the focal slice regions in the presented and the focal sweep methods respectively for the case of  $N = 64$ . These results found that the presented method has almost constant RMSEs around  $1.1 \times 10^{-13}$  over all depths. In contrast, RMSEs of the focus sweep method vary according to depths and is increased at depths far from the middle. This is due to the modeling error of the averaged focal stack image.

### 3.3 Evaluation of Noise Sensitivity

Since the presented filtering process is linear, its noise sensitivity can be quantitatively evaluated when the additive

noise is white. We here define noise amplification factor (NAF) as the metric of noise sensitivity by  $RMSE/\sigma_n$ , where  $RMSE$  represents RMSE of the reconstructed image and  $\sigma_n$  the standard deviation of the additive noise. In this simulation, the NAF of the presented method is given by

$$NAF = \sqrt{\frac{1}{512^2} \sum_{k=0}^{511} \sum_{l=0}^{511} \sum_{i=1}^N |K_i(k, l)|^2}, \quad (16)$$

where  $K_i(k, l)$  denote the discrete Fourier transform of the synthetic filters and  $(k, l)$  the indices of spatial frequencies. Similarly, for the focus sweep method, the NAF can be given by

$$NAF = \sqrt{\frac{1}{512^2 \cdot N} \sum_{k=0}^{511} \sum_{l=0}^{511} |D(k, l)|^2}, \quad (17)$$

where  $D(k, l)$  denotes the discrete Fourier transform of the deconvolution filter. Note that the noise standard deviation is reduced to  $\sigma_n/\sqrt{N}$  because the average image of the focal stack images is computed.

The NAFs computed for both presented and focus sweep methods are shown in Table 2. In the simulation settings, all the NAFs of the presented method are less than 3, which indicates the presented method is not so sensitive to noise as most image recovery methods have in usual. This is because the presented filters are stable and the frequency characteristics have 1 at a maximum as shown in Fig. 2. On the other hand, for  $N = 8, 16, 32$ , the focus sweep method has higher NAFs and more sensitive to noise than the presented method. But, for  $N = 64$ , thanks to the averaging effect, the NAF is decreased and is lower than that of the presented method.

Table 3 shows the NAFs calculated using the reconstructed images from the focal stack images with white Gaussian noise of  $\sigma_n = 10$  and 20. When calculating the NAF, we used the reconstructed images and the focal stack images after unsigned 8-bit quantization. The calculated

NAFs are similar to the theoretical values in Table 2. For the case of  $N = 8$  in the focus sweep method, the NAFs are lower than the theoretical values because of the clipping at 0 and 255 in the quantization.

Figure 4 shows the simulation results using ‘‘Baboon’’ test image for  $N = 8$  and  $\sigma_n = 10$ . As shown in Fig. 4 (a), the focal stack images were generated from the ground truth image (Fig. 4 (b)). Each focal stack image consists of eight regions in different focus settings. For example, the image  $g_1$  is focused on the nearest region (the most left rectangle) and is out of focus in the other regions; the image  $g_8$  is focused on the farthest region (the most right rectangle) and is out of focus in the other regions.

The all-in-focus image reconstructed by the presented filter bank is shown in Fig. 4 (c). It can be seen that the reconstructed image appears in-focus on all eight regions and is close to the ground truth except the region boundaries. The noise in the reconstructed image is amplified by 2.51 times (see Table 3), but this is not severe quality degradation. For comparison, the all-in-focus image reconstructed by the conventional focus sweep method [25], [26] is shown in Fig. 4 (e), which is deconvoluted from the averaged image (Fig. 4 (d)) of the focal stack. In this result, noise is amplified by 8.20 times and the quality is degraded.

The another simulation results using ‘‘Peppers’’ test image for  $N = 64$  and  $\sigma_n = 20$  are shown in Fig. 5. Figure 5 (a) shows examples of the focal stack images generated from the ground truth image in Fig. 5 (b). The generated focal stack images have continuously 64 depth regions with different focus effects. As shown in Fig. 5 (c), the reconstructed image by the presented filter bank looks in-focus over all the regions. The noise is amplified only by 1.67 and the quality is not much degraded compared with that in the focal stack images. Figure 5 (e) shows the all-in-focus image reconstructed by the focus sweep method [25], [26] from the averaged focal stack image (Fig. 5 (d)). Although the focal slices in middle layers are recovered in-focus, the nearest and the farthest regions are slightly blurry compared to the ground truth. This is because PSF of the averaged image is approximately modeled to be space-invariant PSF.

As future work, to suppress noise amplification, noise reduction of the focal stack is needed. Some robust noise reduction methods for video data will be effective, since a set of the focal stack images is a 3D data same as video data.

**Table 2** Noise amplification factors  $RMSE/\sigma_n$ .

method	Number of layers, $N$			
	8	16	32	64
Presented	2.56	2.72	2.32	1.78
focus sweep [25], [26]	10.9	4.05	2.44	1.64

**Table 3** Calculated noise amplification factors using three test images in presence of white Gaussian noise.

test image	method	Standard deviation of Gaussian noise, $\sigma_n$							
		10				20			
		Number of layers, $N$				Number of layers, $N$			
		8	16	32	64	8	16	32	64
Baboon	Presented	2.51	2.67	2.28	1.77	2.37	2.51	2.18	1.72
	Focus sweep [25], [26]	8.20	3.92	2.44	1.69	5.36	3.41	2.30	1.61
Lenna	Presented	2.51	2.66	2.28	1.77	2.36	2.49	2.16	1.70
	Focus sweep [25], [26]	8.27	3.90	2.44	1.69	5.47	3.41	2.29	1.60
Peppers	Presented	2.47	2.62	2.25	1.74	2.30	2.42	2.10	1.67
	Focus sweep [25], [26]	8.22	3.84	2.43	1.72	5.50	3.35	2.24	1.59

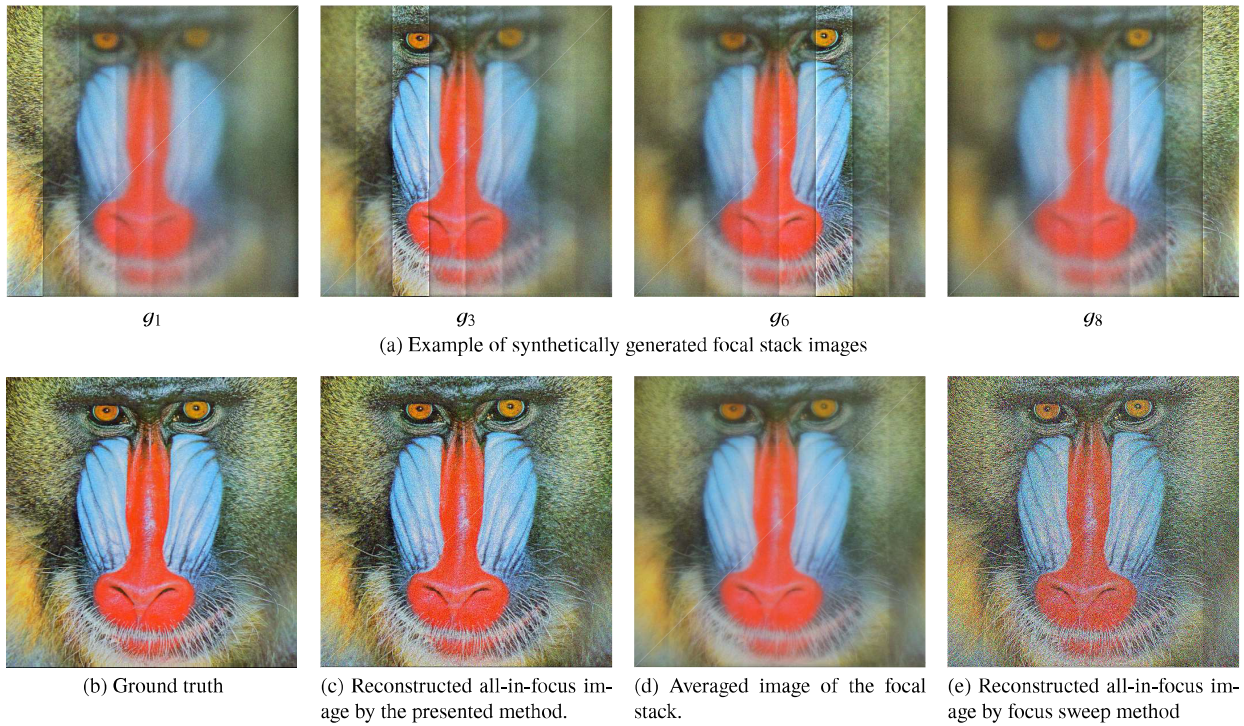


Fig. 4 Simulation results using “Baboon” test image for  $N = 8$  and  $\sigma_n = 10$ .

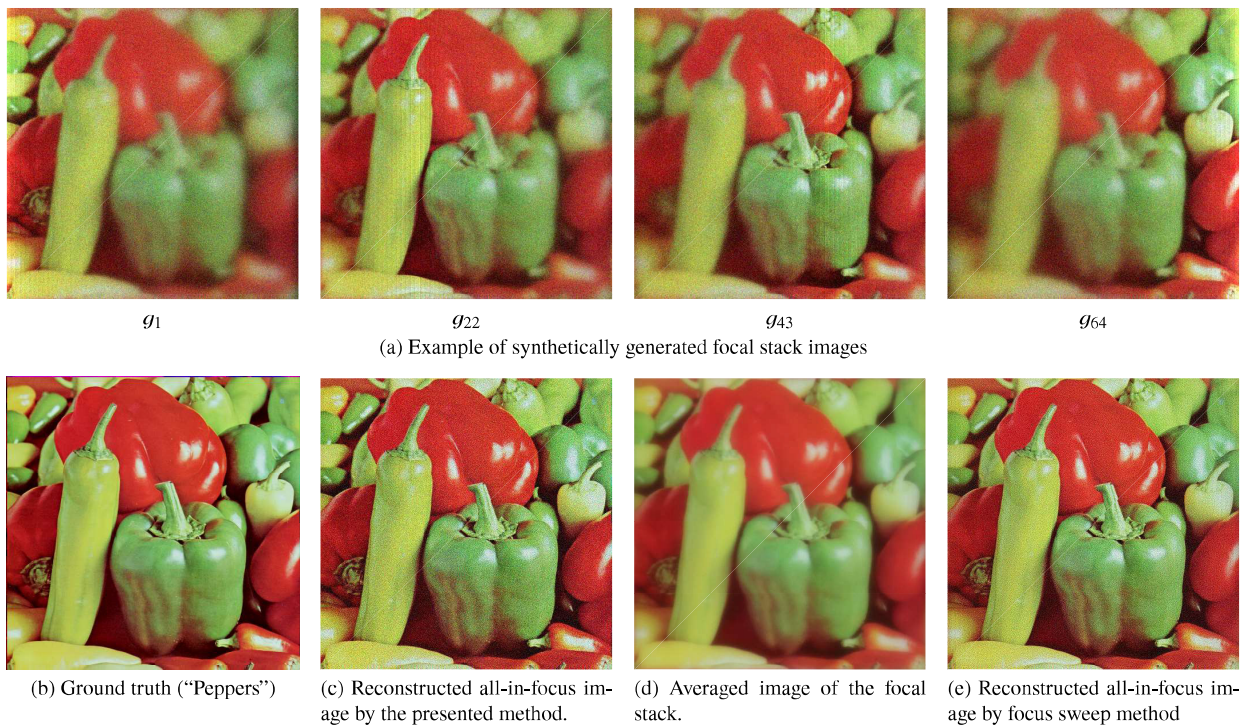


Fig. 5 Simulation results using “Peppers” test image for  $N = 64$  and  $\sigma_n = 20$ .

#### 4. Conclusion

The aperture function was theoretically derived as 2D Cauchy function for perfectly reconstructing all-in-focus

image by filtering the focal stack images. The filters were simply designed using the function. In the simulation, the all-in-focus image was successfully reconstructed with high quality by the presented filters without region segmentation.

In future, we will evaluate the performance of the pre-

sented method through experiments using real captured focal stack images. To do this, we need to design not only aperture but also lens system to create Cauchy PSF as precisely as possible on the captured focal stack. In addition, extending to producing re-focusing and perspective shifting effects will be considered.

## References

- [1] D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE Signal Process. Mag.*, vol.13, no.3, pp.43–64, 1996.
- [2] E.R. Dowski and W. Cathey, "Extended Depth of Field Through Wavefront Coding," *Applied Optics*, vol.34, no.11, pp.1859–1866, 1995.
- [3] S. Bradburn, W.T. Cathey and E.R. Dowski, "Realizations of focus invariance in optical–digital systems with wavefront coding," *Appl. Opt.*, vol.36, no.35, pp.9157–9166, 1997.
- [4] N. George and W. Chi, "Extended Depth of Field Using a Logarithmic Asphere," *J. Optics A: Pure and Applied Optics*, vol.5, pp.157–163, 2003.
- [5] A. Castro and J. Ojeda-Castaneda, "Asymmetric Phase Masks for Extended Depth of Field," *Applied Optics*, vol.43, no.17, pp.3474–3479, 2004.
- [6] G. Häusler, "A Method to Increase the Depth of Focus by Two Step Image Processing," *Optics Comm.*, vol.6, no.1, pp.38–42, 1972.
- [7] S. Kuthirummal, H. Nagahara, C. Zhou, and S.K. Nayar, "Flexible Depth of Field Photography," *IEEE Transactions on Pattern Recognition and Machine Intelligence*, vol.33, no.1, pp.58–71, 2011.
- [8] D. Míau, O. Cossairt, and S.K. Nayar, "Focal sweep videography with deformable optics," *IEEE International Conference on Computational Photography*, pp.1–8, 2013.
- [9] H.A. Eltoukhy and S. Kavusi, "A computationally efficient algorithm for multifocus image reconstruction," *Proc. SPIE Electron. Imaging*, vol.5017, pp.332–341, 2003.
- [10] P.-L. Lin and P.Y. Huang, "Fusion methods based on dynamic-segmented morphological wavelet or cut and paste for multifocus images," *Signal Process.*, vol.88, no.6, pp.1511–1527, 2008.
- [11] S. Li and B. Yang, "Multifocus image fusion using region segmentation and spatial frequency," *Image Vis. Comput.*, vol.26, no.7, pp.971–979, 2008.
- [12] P.J. Burt and R.J. Kolczynski, "Enhanced image capture through fusion," *Proc. Int. Conf. Comput. Vis.*, pp.173–182, 1993.
- [13] B. Forster, D. van de Ville, J. Berent, D. Sage, and M. Unser, "Complex wavelets for extended depth-of-field: A new method for the fusion of multichannel microscopy images," *Microsc. Res. Tech.*, vol.65, no.1-2, pp.33–42, 2004.
- [14] J. Tian and L. Chen, "Multi-focus image fusion using wavelet-domain statistics," *Proc. Int. Conf. Image Process.*, pp.1205–1208, 2010.
- [15] J. Tian, L. Chen, L. Ma, and W. Yu, "Multi-focus image fusion using a bilateral gradient-based sharpness criterion," *Opt. Commun.*, vol.284, no.1, pp.80–87, 2011.
- [16] B. Yang and S. Li, "Multifocus image fusion and restoration with sparse representation," *IEEE Trans. Instrum. Meas.*, vol.59, no.4, pp.884–892, 2010.
- [17] W. Huang and Z. Jing, "Evaluation of focus measures in multifocus image fusion," *Pattern Recognit. Lett.*, vol.28, no.4, pp.493–500, 2007.
- [18] A. Kubota and K. Aizawa, "Reconstructing arbitrarily focused images from two differently focused images using linear filters," *IEEE Trans. Image Process.*, vol.14, no.11, pp.1848–1859, 2005.
- [19] K. Kodama, K. Aizawa, and M. Hatori, "Generation of arbitrarily focused images by using multiple differently focused images," *J. Elect. Imag.*, vol.7, no.1, pp.138–144, 1998.
- [20] K. Aizawa, K. Kodama, and A. Kubota, "Producing object based special effects by fusing multiple differently focused images," *IEEE Trans. Circuits Syst. Video Technol.*, vol.10, no.2, pp.323–330, 2000.
- [21] J.R. Alonso, A. Fernández, and J.A. Ferrari, "Reconstruction of perspective shifts and refocusing of a three-dimensional scene from a multi-focus image stack," *Applied Optics*, vol.55, no.9, pp.2380–2386, 2016.
- [22] K. Kodama, H. Mo, and A. Kubota, "Free viewpoint iris and focus image generation by using a three-dimensional filtering based on frequency analysis of blurs," *Proc. Int. Conf. Acoust. Speech Signal Process.*, vol.2, pp.625–628, 2006.
- [23] K. Kodama and A. Kubota, "Free iris and focus image generation by merging multiple differently focused images," *Trans. Inf. Syst.*, vol.E90-D, no.1, pp.191–198, 2007.
- [24] K. Kodama, H. Mo, and A. Kubota, "Simple and fast all-in-focus image reconstruction based on three-dimensional/two-dimensional transform and filtering," *Proc. Int. Conf. Acoust. Speech Signal Process.*, vol.1, pp.769–772, 2007.
- [25] K. Kodama and A. Kubota, "Efficient Reconstruction of All-in-Focus Images Through Shifted Pinholes from Multi-Focus Images for Dense Light Field Synthesis and Rendering," *IEEE Transactions on Image Processing*, vol.22, no.11, pp.4407–4421, 2013.
- [26] A. Levin and F. Durand, "Linear view synthesis using a dimensionality gap light field prior," *IEEE Conference on Computer Vision and Pattern Recognition*, pp.1831–1838, 2010.
- [27] C. Zhou, S. Lin, and S. Nayar, "Coded aperture pairs for depth from defocus," *Proc. International Conference on Computer Vision*, pp.325–332, 2009.
- [28] H. Farid and E.P. Simoncelli, "Range estimation by optical differentiation," *Journal of the Optical Society of America A*, vol.15, no.7, pp.1777–1786, 1998.
- [29] S. Hiura and T. Matsuyama, "Depth measurement by the multi-focus camera," *IEEE Conference on Computer Vision and Pattern Recognition*, pp.953–959, 1998.
- [30] A. Kubota, "Synthesis filter bank and pupil function for perfect reconstruction of all-in-focus image from focal stack," *Proc. SPIE Int. Conf. Quality Control by Artificial Vision*, 2017.
- [31] Y.Y. Schechner, N. Kiryati, and R. Basri, "Separation of Transparent Layers using Focus," *International Journal of Computer Vision*, vol.39, no.1, pp.25–39, 2000.
- [32] U. Grenander and G. Szegő, "Toeplitz forms and their applications," *Uni. Calif. Press*, vol.11, no.10, p.38, 1958.
- [33] S. Kotz and S. Nadarajah, "Multivariate t-Distributions and Their Applications," *Cambridge Univ. Press*, pp.87–126, 2004.



**Akira Kubota** received the B.E. degree in electrical engineering from Oita University, Oita, Japan, in 1997, and the M.E. and Dr. E. degrees in electrical engineering from the University of Tokyo, Tokyo, Japan, in 1999 and 2002, respectively. He is currently an Associate Professor with the Department of Electrical, Electronic, and Communication Engineering, Chuo University, Tokyo. From September 2003 to July 2004, he was with the Advanced Multimedia Processing Laboratory, Carnegie Mellon University, Pittsburgh, PA, USA, as a Research Associate. He was a Research Fellow with the Japan Society for the Promotion of Science from April 2002 to July 2004. His current research interests include image coding, image reconstruction and computational photography.



**Kazuya Kodama** received the B.E., M.E., and the Dr. E. degrees in electrical engineering from the University of Tokyo, Tokyo, Japan, in 1994, 1996, and 1999, respectively. He is currently an Associate Professor with the National Institute of Informatics, Research Organization of Information and Systems, Tokyo. He is an Associate Professor with the Department of Informatics, Graduate University for Advanced Studies, Kanagawa, Japan. From April 1998 to March 1999, he was a Research Fellow with the

Japan Society for the Promotion of Science. His current research interests include image acquisition, image representation, image reconstruction, and 3-D image coding.



**Asami Ito** received the B.E. degree in electrical, electronic and communication engineering from Chuo University, Tokyo, Japan, in 2018. She is currently a master student with school of science and engineering, Chuo University. Her current research interests include image reconstruction and computational photography.