# 3D-Sensing Technology based on Lensless Light-Field Imaging with Fresnel Zone Aperture

Keita Yamaguchi, Yusuke Nakamura, Kazuyuki Tajima, Toshiki Ishii,
Koji Yamasaki, Takeshi Shimano
Center for Technology Innovation – Electronics, Hitachi, Ltd.,
292 Yoshida, Totsuka, Yokohama, 244-0817 Japan
E-mail:keita.yamaguchi.oe@hitachi.co.jp

**Abstract**　Lensless light-field imaging technology with Fresnel zone aperture reconstructs a complex image, composed of complex numbers, from captured sensor images. In the present study, we found that the imaginary part of the complex image indicates zero in the vicinity of the in-focus position. By using this characteristic, we established a new 3D-sensing algorithm which generates depth images without light sources or binocular vision. The algorithm enables compact 3D-sensors suitable for various applications, e.g. smartphones, robots, and vehicles.

**Keywords:** lensless, light-field, Fresnel zone, depth, 3D, computational photography

## 1. Introduction

The progress in image sensor technology have enabled cameras to become smaller and smaller. Recently, lens has become the stumbling block for further downsizing of the traditional cameras. On the other hand, progress in signal processing technology has lead emergence of technologies called "computational photography." Within this emerging field of technologies, lensless camera appears to be the technology which enables further downsizing of the cameras [1].

Previously, we reported lensless light-field imaging with a Fresnel zone aperture (FZA). In addition to the thin form factor common in the lensless cameras, the proposed technology possesses beneficial features such as light calculation load in the image reconstruction processing and refocusing, i.e. focus adjustment after capturing images [2]. The ability to refocus indicates that the proposed lensless camera records depth information of the scene. For this reason, we aimed to exploit this feature to establish a new 3D-sensing algorithm which generates depth image from the captured sensor image. Since the proposed method allows 3D-sensing without using special light sources or binocular vision, compact 3D-sensors can be developed. Moreover, the non-existence of the special light sources enables usage of multiple numbers of the 3D-sensors without having have to prevent interference among the sensors. We expect that the 3D-sensors with these features are suitable for applications in various fields, e.g. smartphones, robots, and vehicles.
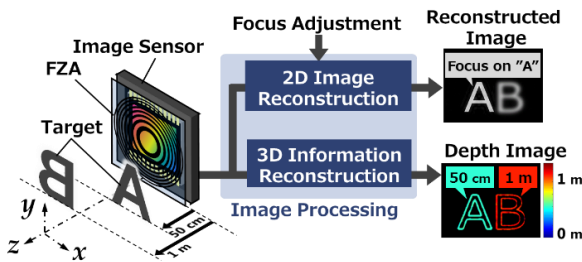
## 2. Reconstruction method



Fig. 1. System overview of the developed lensless camera.

Figure 1 shows the overview of the lensless camera with FZA. The optical part of this system consists of an image sensor and a FZA. The FZA is placed several millimeters apart from the sensor surface. The scene in front of the lensless camera casts shadow of the FZA onto the sensor. The sensor captures this shadow as a sensor image. The scene can be reconstructed by image processing of this sensor image.

The basic principle of capturing and reconstructing 2D images in our lensless camera is as follows. In order to apply noise cancelling by the fringe scan method [2], the lensless camera captures four sensor images with four different FZAs. The transmissivity of the four patterns are described by

$$T_N(r) = \frac{1}{2}[1 + \cos(\beta r^2 + \phi_N)], \qquad (1)$$

$$\phi_N = \phi_0 + \frac{N\pi}{2}, \qquad (2)$$

for $N = 0, 1, 2, 3$, where $r$ is a radial coordinate in the FZA plane, $\beta$ is a coefficient which defines the pitch of the FZAs, and $\phi_0$ is the initial phase of the FZAs. In the actual prototype lensless cameras, $T_N(r)$ in Eq. (1) are approximated by binary masks. Sensor images with four different FZAs can be obtained either by time division method or spatial division method. Time division method uses liquid crystal devices to change the transmissivity $T_N(r)$ over time. On the other hand, spatial division method tiles up the four different FZAs in front of a single image sensor.

A scene at a distance $d_1$ away from the FZA surface casts shadow with the intensity distribution $I_N(x, y)$ onto the sensor.

$$I_N(x, y) = f(x, y) * \frac{1}{2}[1 + \cos(\beta' r^2 + \phi_N)], \qquad (3)$$

$$\beta' = \frac{d_1^2}{(d_1 + d_2)^2}\beta. \qquad (4)$$

Here, $f(x, y)$ is the intensity distribution of the scene at distance $d_1$, $d_2$ is the distance between the FZA and the image sensor, and * denotes convolution operator.

In order to reconstruct the 2D image of the scene, the four sensor images are unified into a complex sensor image as shown in Eq. (5).

$$\begin{aligned} I_{\text{Comp}}(x, y) &= \sum_{N=0}^{3} I_N(x, y)\, e^{-i\phi_N} \\ &= f(x, y) * \exp[i(\beta' r^2 + \phi_0)]. \end{aligned} \qquad (5)$$

Then, the complex sensor image is cross correlated with a virtual FZA (a FZA generated in a computer) which is described by

$$I_{\text{VFZA}}(x, y) = \exp[i(\beta'' r^2 + \phi_0)], \qquad (6)$$

where $\beta''$ is the pitch coefficient of the virtual FZA. When $\beta'' = \beta'$, the cross correlation $g(x, y)$ yields the 2D image of the

scene as shown in Eq. (7).

$$g(x, y) = I_{\text{Comp}}(x, y) \star I_{\text{VFZA}}(x, y)$$
$$= f(x, y) * \delta(x, y) \qquad (7)$$
$$= f(x, y).$$

Here, $\star$ denotes the cross correlation operator, and $\delta(x, y)$ is the Dirac's delta.

### 3. 3D Sensing algorithm & Experimental results

Next, we will explain the proposed 3D-sensing algorithm. In order to extract the depth information from the sensor image, a virtual FZA whose pitch coefficient $\beta'' = \beta'$ must be found out. For simplicity, we will consider a simple scene which only contains a point light source at $x = 0$, $y = 0$, and distance $d_1$. In this case, the reconstructed image for arbitrary $\beta''$ will be

$$g(x, y) = I_{\text{Comp}}(x, y) \star I_{\text{VFZA}}(x, y)$$
$$= \cos(\beta' r^2 + \phi_0) * \cos(\beta'' r^2 + \phi_0)$$
$$+ \sin(\beta' r^2 + \phi_0) * \sin(\beta'' r^2 + \phi_0) \qquad (8)$$
$$+ i[\sin(\beta' r^2 + \phi_0) * \cos(\beta'' r^2 + \phi_0)$$
$$- \cos(\beta' r^2 + \phi_0) * \sin(\beta'' r^2 + \phi_0)].$$

From Eq. (8), it is observed that the imaginary part of the cross correlation becomes zero when $\beta'' = \beta'$ and becomes non-zero elsewhere. We confirmed this behavior through simulation for $d_1 = 50.0$ mm, $d_2 = 5.5$ mm, $\beta = 6.34$ rad/mm$^2$, and $\phi_0 = 0$. The sensor size and the number of pixels used for the simulation were $11.264 \times 11.264$ mm$^2$ and $1024 \times 1024$, respectively. The result is shown in Fig. 2.
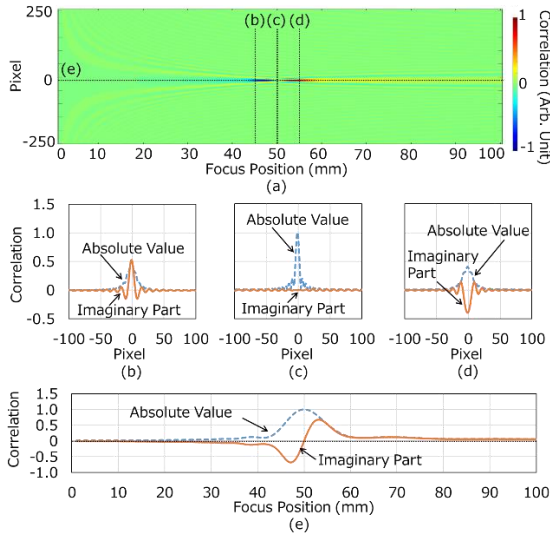


Fig. 2. (a) Map of the imaginary part of $g(x, y)$ along $y = 0$. (b) - (e) The absolute value and the imaginary part of $g(x, y)$ along dotted lines denoted as (b) - (e) in (a). Values in each graphs are normalized with the absolute value of $g(0,0)$ at $\beta'' = \beta'$.

Figure 2 (a) maps the imaginary part of the simulated $g(x, 0)$ for different focus position. The figure shows that as the focus position moves toward the in-focus position, the cross correlation converges into a spot. In addition, the sign of the imaginary part reverses when the focus position crosses the in-focus position. This sign reversal is clearly observed in Figs. 2 (b), (c), and (d) where $g(x, 0)$ for focus at 45 mm, 50 mm, and 55 mm are plotted, respectively. Moreover, as Fig. 2 (e) shows, the imaginary part of $g(0, 0)$ indicates zero at the in-focus position. For this reason, we expected the imaginary part of the cross

correlation to be a good indicator for extracting the depth information. Thus, our 3D-sensing algorithm searches for $\beta''$ where the imaginary part indicates zero, and calculate the depth of the target according to Eq. (4). Since this can be achieved through linear interpolation along the $z$ axis, depth information can be extracted without performing an exhaustive search. This speeds up the computation required for generating depth images.

To verify the effectiveness of the above method, we developed a prototype lensless camera. The sensor size, $d_2$, $\beta$, and $\phi_0$ were the same as those used in the simulation shown in Fig. 2. The prototype performs fringe scan through time division method. The experimental results for various objects placed between 100 mm and 1000 mm are shown in Fig. 3. The figures show that the developed prototype successfully extracts 3D information from the captured sensor images. We also confirmed that within this range, the depth errors are in the order of several percent.
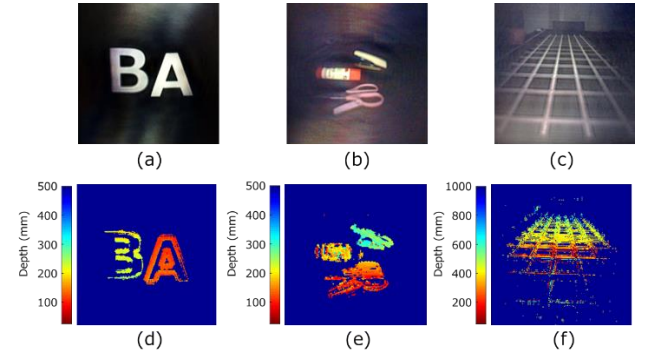


Fig. 3. (a) Image of alphabet objects reconstructed at 180 mm. A and B are placed at different distances. (b) Image of stationaries reconstructed at 200 mm. (c) Image of grid patterns reconstructed at 500 mm. (d) - (f) Depth maps corresponding to (a), (b), and (c), respectively.

### 4. Conclusion

We proposed a new 3D-sensing algorithm for the lensless camera with Fresnel zone aperture. The algorithm exploits the feature of the lensless camera that the imaginary part of the complex image indicates zero in the vicinity of the in-focus position. Through experiment with the developed prototype, we have successfully verified that the proposed algorithm is effective. Since the proposed method enables 3D-sensing without using special light sources or binocular vision, we expect lensless camera based 3D-sensor to be applied in various fields such as smartphones, robots, and vehicles.

In the present method, coherent blur inherent to lensless cameras are one of the main cause of the depth error. We expect suppression of this effect will enhance the depth precision. As a future work, we aim to improve the depth precision in the lensless camera based 3D-sensors.

### References

[1] M. Salman Asif, et.al., "FlatCam: Thin, Lensless Cameras Using Coded Aperture and Computation", IEEE Tran. Comp. Imag., Vol 3, No. 3, pp. 384-397 (2017).

[2] T. Shimano, et al., "Lensless light-field imaging with Fresnel zone aperture: quasi-coherent coding", Appl. Opt., Vol. 57, No. 11, pp. 2841-2850 (2018).