Peer-reviewed version available at Sensors 2019, 19, 1329; doi:10.3390/s19061329

Article

Super Field-of-View Lensless Camera by Coded Image Sensors

Tomoya Nakamura [1,2,*], Keiichiro Kagawa [3], Shiho Torashima [1] and Masahiro Yamaguchi [1]

1 School of Engineering, Tokyo Institute of Technology, Kanagawa 2268502, Japan; nakamura.t.bj@m.titech.ac.jp (T.N.); torashima.s.aa@m.titech.ac.jp (S.T.); yamaguchi.m.aa@m.titech.ac.jp (M.Y.)
2 PRESTO, Japan Science and Technology Agency, Saitama 332-0012, Japan; nakamura.t.bj@m.titech.ac.jp
3 Research Institute of Electronics, Shizuoka University, Shizuoka 432-8011, Japan; kagawa@idl.rie.shizuoka.ac.jp
* Correspondence: nakamura.t.bj@m.titech.ac.jp; Tel.: +81-45-924-5296
† This paper is an extended version of our paper published in Proceedings of the 4th International Workshop on Image Sensors and Imaging Systems, Tokyo, Japan, 28–29 November 2018.

Abstract: A lensless camera is an ultra-thin computational-imaging system. Existing lensless cameras are based on the axial arrangement of an image sensor and a coding mask, and therefore, the back side of the image sensor cannot be captured. In this paper, we propose a lensless camera with a novel design that can capture the front and back sides simultaneously. The proposed camera is composed of multiple *coded image sensors*, which are complementary-metal-oxide-semiconductor (CMOS) image sensors in which air holes are randomly made at some pixels by drilling processing. When the sensors are placed facing each other, the object-side sensor works as a coding mask and the other works as a sparsified image sensor. The captured image is a sparse coded image, which can be decoded computationally by using compressive-sensing-based image reconstruction. We verified the feasibility of the proposed lensless camera by simulations and experiments. The proposed thin lensless camera realizes super field-of-view imaging without lenses or coding masks, and therefore can be used for rich information sensing in confined spaces. This work also suggests a new direction in the design of CMOS image sensors in the era of computational imaging.

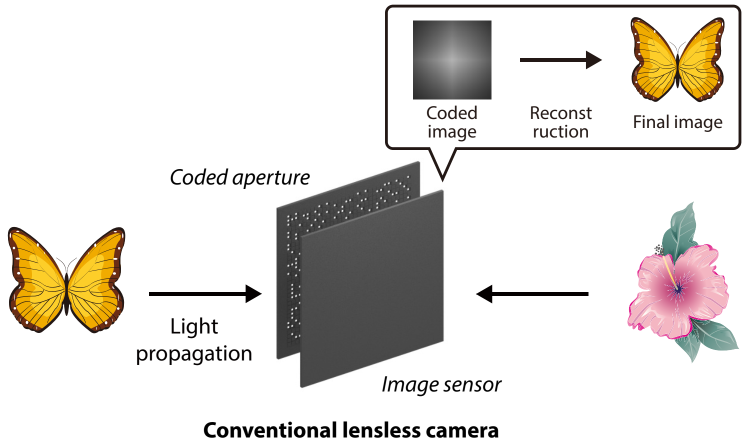
Keywords: computational imaging; lensless camera; CMOS image sensor; compressive sensing

---

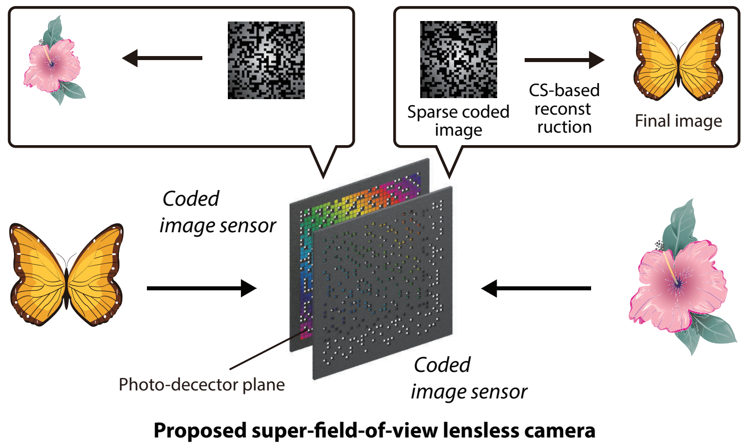
1. Introduction

A computational lensless camera [1] works on the basis of computational imaging, which is a combination of optical encoding and computational decoding [2], instead of lens-based optical imaging. This frees the camera from the need for optical focusing, allowing the camera to be implemented with ultra-thin, miniature hardware [3]. In a lensless camera, a lens system is typically replaced with a coded aperture such as a coding mask, which makes the inverse problem numerically invertible. As the coded aperture, amplitude masks [4–6] or phase-modulation optics [7–10] have been installed in front of the image sensor. In computational lensless imaging, spatial information of a subject is coded by the coded aperture, and the optically coded image is sampled by the image sensor. After exposure, the captured coded image is numerically decoded into the original image of the subject. In the decoding, the inverse problem of coded sensing is solved by signal processing.

Existing architectures of lensless cameras are based on the axial arrangement of an image sensor and a coded aperture, like that shown in Fig. 1(a). This means that the back side of the sensor cannot be captured. In other words, the field-of-view (FOV) in mask-based lensless imaging is limited to 180 degree in the forward direction only. To extend the limit, the use of multiple lensless cameras is the most straightforward



(a)



(b)

Figure 1. Concept of (a) conventional lensless camera and (b) proposed super-field-of-view lensless camera using coded image sensors (CISs). The photo-detector plane faces inside the optical system in both figures.

approach; however, this is naturally at the sacrifice of the simplicity and thinness of the system because additional lensless cameras are required.

In this paper, we propose a lensless camera with a novel design that is capable of capturing images of the front and back sides at the same time. Figure 1(b) shows the concept of the proposed super-FOV lensless camera. The proposed camera is composed of an axial arrangement of two *coded image sensors* (CISs), which are complementary-metal-oxide-semiconductor (CMOS) image sensors in which air holes are randomly formed at some pixels by drilling processing. In this arrangement, the photo-detector planes of the sensors face each other. Thanks to the drilling-based physical sparsification of the image sensor, the sensor works not only as an image sensor but also as a coded aperture for the opposing image sensor. After image capturing, the captured sparse coded images are decoded into the original images of the subjects by a numerical image-reconstruction algorithm based on the compressive-sensing (CS) framework [11–13]. Section 2 describes the methodology of the proposed lensless imaging, Sec. 3 shows the concept verification and performance analysis by numerical simulations, and Sec. 4 shows the experimental verification using a real mask-based lensless camera and simulated air holes. This paper is an extended version of our earlier conference paper on this camera system [14].

Peer-reviewed version available at Sensors 2019, 19, 1329; doi:10.3390/s19061329

2. Method

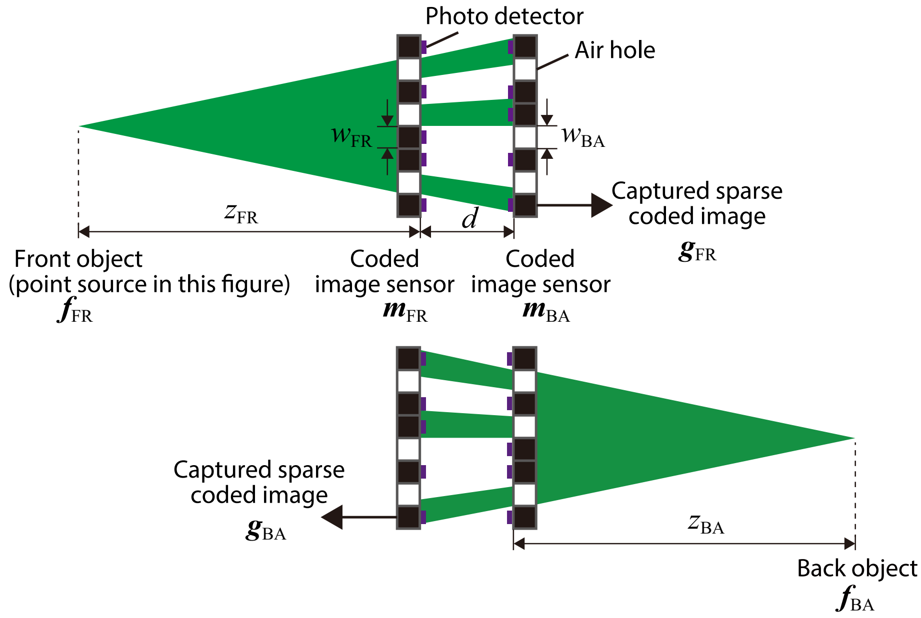


Figure 2. Parameter definition for the proposed camera. The illustration at each image sensor represents the amplitude transmittance of the pixel area, where white pixels correspond to air holes. The green area represents the light transmittance in an optical system when the object is a point source.

Figure 2 presents a schematic view of the optical hardware of the proposed camera, with definitions of the system parameters. As the optical hardware, two CISs that include both photo-detectors and air holes are provided. The photo-detector planes face the inside of the optical system. We assume that the air holes in the image sensor are implemented by drilling processing, where the drilled pixels are chosen randomly. In the set-up, we call the object at the left side in figure the front object $\boldsymbol{f}_{\mathrm{FR}} \in \mathbb{R}^{N_{\mathrm{f}}\times 1}$, and the object at the right side the back object $\boldsymbol{f}_{\mathrm{BA}} \in \mathbb{R}^{N_{\mathrm{f}}\times 1}$. Here, each vector represents the light intensity of an object. Similarly, we denote the front CIS as $\boldsymbol{m}_{\mathrm{FR}} \in \mathbb{R}^{N_{\mathrm{f}}\times 1}$ and the back CIS as $\boldsymbol{m}_{\mathrm{BA}} \in \mathbb{R}^{N_{\mathrm{f}}\times 1}$. Note that the values of all the elements in a sensor vector are zero or one, which physically expresses the amplitude transmittance of the pixels. In addition, we also denote the intensity vectors of captured sparse coded images as $\boldsymbol{g}_{\mathrm{BA}} \in \mathbb{R}^{N_{\mathrm{g}}\times 1}$ and $\boldsymbol{g}_{\mathrm{FR}} \in \mathbb{R}^{N_{\mathrm{g}}\times 1}$, respectively. The sizes of the square air holes in each sensor are denoted as w_{FR} and w_{BA}. The axial interval between the two sensors is denoted as d, and the distances between the sensors and the objects are denotes as z_{FR} and z_{BA}, respectively.

The forward model of the coded image sensing by the proposed camera is as follows:

$$\boldsymbol{g}_{\mathrm{FR}} = \boldsymbol{M}_{\mathrm{BA}} \left(\boldsymbol{h}_{\mathrm{FR}} * \boldsymbol{f}_{\mathrm{FR}} \right), \tag{1}$$

$$\boldsymbol{g}_{\mathrm{BA}} = \boldsymbol{M}_{\mathrm{FR}} \left(\boldsymbol{h}_{\mathrm{BA}} * \boldsymbol{f}_{\mathrm{BA}} \right), \tag{2}$$

where $\boldsymbol{M}_{\mathrm{BA}} \in \mathbb{R}^{N_{\mathrm{g}}\times N_{\mathrm{f}}}$ and $\boldsymbol{M}_{\mathrm{FR}} \in \mathbb{R}^{N_{\mathrm{g}}\times N_{\mathrm{f}}}$ are the matrices expressing the sparse sampling by CISs whose structures are $\boldsymbol{m}_{\mathrm{BA}}$ and $\boldsymbol{m}_{\mathrm{FR}}$; and $\boldsymbol{h}_{\mathrm{FR}} \in \mathbb{R}^{N_{\mathrm{f}}\times 1}$ and $\boldsymbol{h}_{\mathrm{BA}} \in \mathbb{R}^{N_{\mathrm{f}}\times 1}$ are the vectors of optical point-spread functions (PSFs), which are the optical shadows of the opposing CISs. The optical PSFs are modeled as follows:

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 February 2019 doi:10.20944/preprints201902.0265.v1

Peer-reviewed version available at Sensors 2019, 19, 1329; doi:10.3390/s19061329

$$\boldsymbol{h}_{\mathrm{FR}} = \boldsymbol{W}_{\mathrm{zFR}}\boldsymbol{m}_{\mathrm{FR}} * \boldsymbol{b}, \tag{3}$$
$$\boldsymbol{h}_{\mathrm{BA}} = \boldsymbol{W}_{\mathrm{zBA}}\boldsymbol{m}_{\mathrm{BA}} * \boldsymbol{b}, \tag{4}$$

where $\boldsymbol{W}_{\mathrm{zFR}} \in \mathbb{R}^{N_{\mathrm{f}} \times N_{\mathrm{f}}}$ and $\boldsymbol{W}_{\mathrm{zBA}} \in \mathbb{R}^{N_{\mathrm{f}} \times N_{\mathrm{f}}}$ are the matrices expressing the scaling of geometrical shadows depending on the distance to the object, and $\boldsymbol{b} \in \mathbb{R}^{N_{\mathrm{f}} \times 1}$ is the blur kernel of diffraction caused by the light propagation between the two image sensors. Here, $*$ denotes convolution. We assume that the optical PSF is approximately space-invariant [10].

In imaging, the patterns of the optical PSF $\boldsymbol{h}$ and the sparse sampling $\boldsymbol{W}$ can be determined by calibration, and $\boldsymbol{g}$ can be obtained by image capturing. Basically, the intensity image of the original object $\boldsymbol{f}$ can be numerically reconstructed by solving the inverse problem of the forward model in Eqs. (1)–(2). However, the inverse problems are ill-posed because of the sparse sampling by the CISs. In other words, $N_{\mathrm{g}} < N_{\mathrm{f}}$ in the inverse problem of the proposed camera. Therefore, we adopt the CS framework [11,12] for image decoding [13]. In the CS framework, the original image can be reconstructed uniquely even when $N_{\mathrm{g}} < N_{\mathrm{f}}$ by solving the following problem:

$$\hat{\boldsymbol{f}}_{\mathrm{FR}} = \underset{\boldsymbol{f}_{\mathrm{FR}}}{\operatorname{argmin}} ||\boldsymbol{g}_{\mathrm{FR}} - \boldsymbol{M}_{\mathrm{BA}}\left(\boldsymbol{h}_{\mathrm{FR}} * \boldsymbol{f}_{\mathrm{FR}}\right)||_{\ell_2} + \tau||\Phi(\boldsymbol{f}_{\mathrm{FR}})||_{\ell_1}, \tag{5}$$
$$\hat{\boldsymbol{f}}_{\mathrm{BA}} = \underset{\boldsymbol{f}_{\mathrm{BA}}}{\operatorname{argmin}} ||\boldsymbol{g}_{\mathrm{BA}} - \boldsymbol{M}_{\mathrm{FR}}\left(\boldsymbol{h}_{\mathrm{BA}} * \boldsymbol{f}_{\mathrm{BA}}\right)||_{\ell_2} + \tau||\Phi(\boldsymbol{f}_{\mathrm{BA}})||_{\ell_1}, \tag{6}$$

where $||\cdot||_{\ell_2}$ and $||\cdot||_{\ell_1}$ denotes the ℓ_2-norm and ℓ_1-norm of a vector, and τ is a constant. $\Phi(\cdot)$ is a regularizer, which corresponds to a linear transformation for sparse modeling, such as the discrete wavelet transform, the discrete cosine transform, and so on. In this paper, we specifically use the two-dimensional (2D) total variation (TV) [15] as the regularizer, which is computed as follows:

$$||\Phi_{\mathrm{TV}}(\boldsymbol{f})||_{\ell_1} = \sum_{n_{\mathrm{x}}}\sum_{n_{\mathrm{y}}} |\nabla \boldsymbol{f}_{\mathrm{x,y}}|, \tag{7}$$

where $\nabla \boldsymbol{f}_{\mathrm{x,y}}$ is the differential images of $\boldsymbol{f}$ along to horizontal and vertical directions, and n_{x} and n_{y} are the pixel count along to the two axes of the 2D image. To implement the well-conditioned inverse problem, the pattern of drilling on image sensors are designed as random patterns [16].

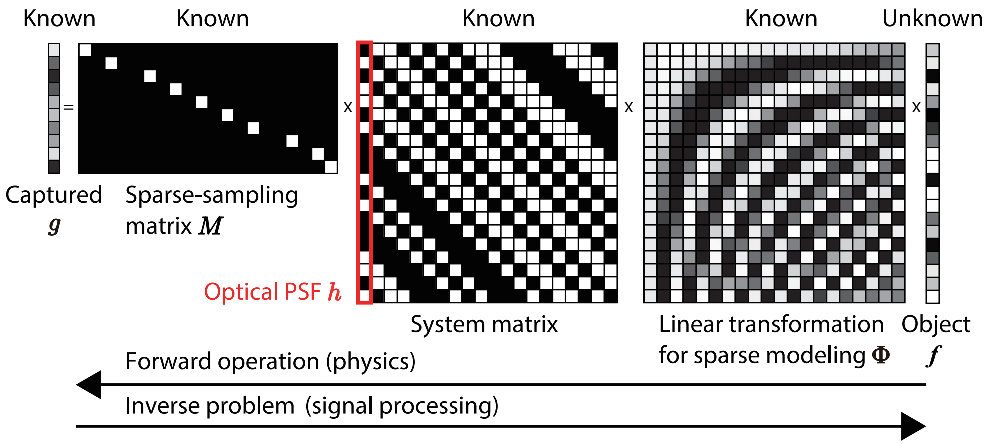


Figure 3. Matrix-vector expression of the imaging model with a one-dimensional (1D) object.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 February 2019 doi:10.20944/preprints201902.0265.v1

Peer-reviewed version available at *Sensors* **2019**, *19*, 1329; doi:10.3390/s19061329

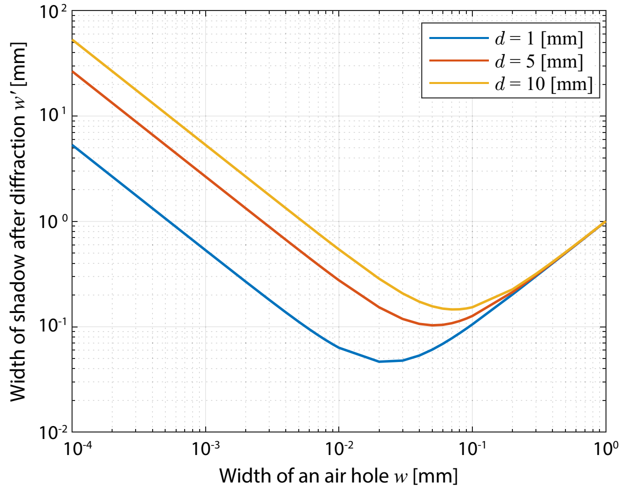


Figure 4. The size of the optical shadow of a small air hole in a CIS, which is formed on another CIS after diffraction. The color of the lines expresses the axial distance between the two image sensors, ranging from 1 mm to 10 mm.

For a more intuitive understanding, the matrix-vector expression of the 1D imaging model in the proposed method is shown in Fig. 3. As mentioned above, we assume that the object after a linear transformation $\Phi(\boldsymbol{f}_{\mathrm{FR}})$ is sparse. Since the optical image on the image plane is the convolution of the mask and the object, the optical system matrix is the Teplitz matrix whose column represents the optical PSF. By the sampling $\boldsymbol{M}_{\mathrm{BA}}$ of a CIS, the length of the captured coded image is made shorter than that of the object. Therefore, the inverse problem of the linear forward imaging model is ill-posed. To solve this, a technique based on CS is necessary.

In practice, the blur kernel $\boldsymbol{b}$ can be directly measured; however, it can also be analyzed by a diffraction calculation [17]. The 1D diameter of the kernel b_{w} can be roughly estimated by using a simple model based on the Fraunhofer approximation as follows:

$$b_{\mathrm{w}} \approx \frac{\lambda d}{w}, \tag{8}$$

where λ is wavelength, and w is the width of an air hole. In this paper, we denote the diameter b_{w} by the distance between first-zero values in the diffraction pattern expressed by the sinc function. Considering diffraction, the size of the optical shadow of the air hole on the image plane is simply formulated as follows:

$$w' = w + b_{\mathrm{w}}. \tag{9}$$

Example values of w' at a wavelength of 532 nm are plotted in Fig. 4. As shown in the figure, there are optimal sizes of the air hole to give the smallest image-spot size after diffraction. The size of the optical shadow of a hole directly corresponds to the highest spatial optical resolution in aperture-shadow-based lensless imaging. Note that the spatial resolution of the lensless camera, independently of the optical resolution, is also restricted by the spatial sampling resolution, which is physically determined by the pixel pitch.

Peer-reviewed version available at Sensors 2019, 19, 1329; doi:10.3390/s19061329

3. Simulations with numerical data

We performed simulations with numerically-generated data to validate the proposed imaging model and its performance. In this paper, we assumed an optical system like that in Fig. 2, and only simulated 1D imaging, as shown in the upper illustration, for simplicity; however, the same discussion can be adopted for imaging in another direction, as shown in the lower illustration. For explanation, here we refer to the captured object as the front object and the object-side image sensor as the front sensor. In this simulation, we generated captured data by emulating imaging and decoded it via a reconstruction algorithm.

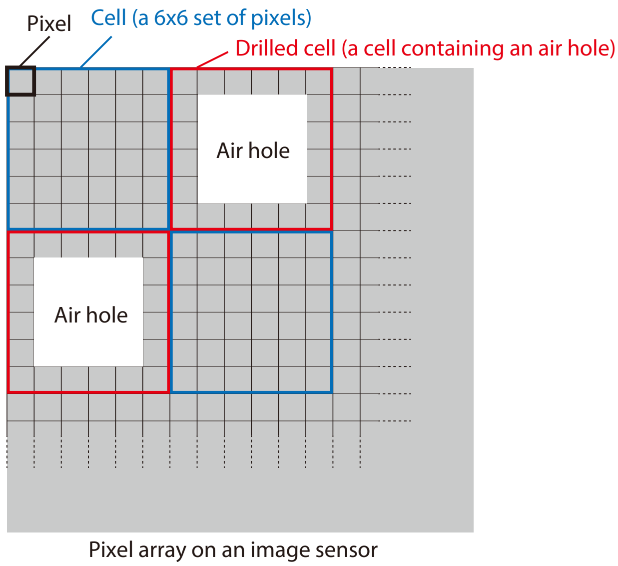


Figure 5. Definition of *cell* and its relation with air holes in a pixel array on an image sensor.

Table 1. Specifications for simulation.

Parameter	Value
Pixel count of object and reconstructed image	384×384 [pixels]
Pixel count of original image sensor (before drilling)	384×384 [pixels]
Pixel pitch in image sensor	10 [μm]
Size of cell in image sensor	60 [μm] $\times$ 60 [μm]
Size of elemental hole in drilled cell	40 [μm] $\times$ 40 [μm]
Thickness of each sensor device	1.0 [mm]
Distance between sensors	2.0 [mm]
Bit depth of signal from pixel	12 [bit]
Wavelength of light	532 [nm]

Here we define the new term *cell* in a pixel array as a set of 6×6 pixels, as shown in Fig. 5. We randomly classified the cells into normal cells and drilled cells, where each drilled cell contained an air hole at the center, as in the figure. For the discussion in this paper, we define the term *air-hole ratio* as the ratio of drilled cells to normal cells. Parameters of the optical system, including the image sensors, are summarized in Table 1. With these parameters, the image sensor can be regarded as an array of 64×64 cells, whereas the total pixel count in the image sensor was 384×384.

As the procedure for the simulation, first, the coded image formed on the back sensor was numerically generated by the convolution of the subject and the optical PSF. The optical PSF was calculated by the convolution of the front sensor and the diffraction-based blur kernel. Note that here we modeled the blur

Peer-reviewed version available at Sensors 2019, 19, 1329; doi:10.3390/s19061329

kernel (***b*** in Eq. (3)) as a 2D Gaussian profile whose full-width at half maximum (FWHM) was determined by that in the Fraunhofer-diffraction pattern of a small air hole (b_{w} in Eq. (8)). Afterward, the digital sampling of the coded image formed by a CIS was emulated with image sparsification, multiplying by the photon-shot noise, and 12-bit quantification. Finally, the captured data were decoded into an original image of the subject using a CS-based image-reconstruction algorithm. For the image reconstruction, we used the TwIST algorithm [18], which iteratively solves Eq. (5).

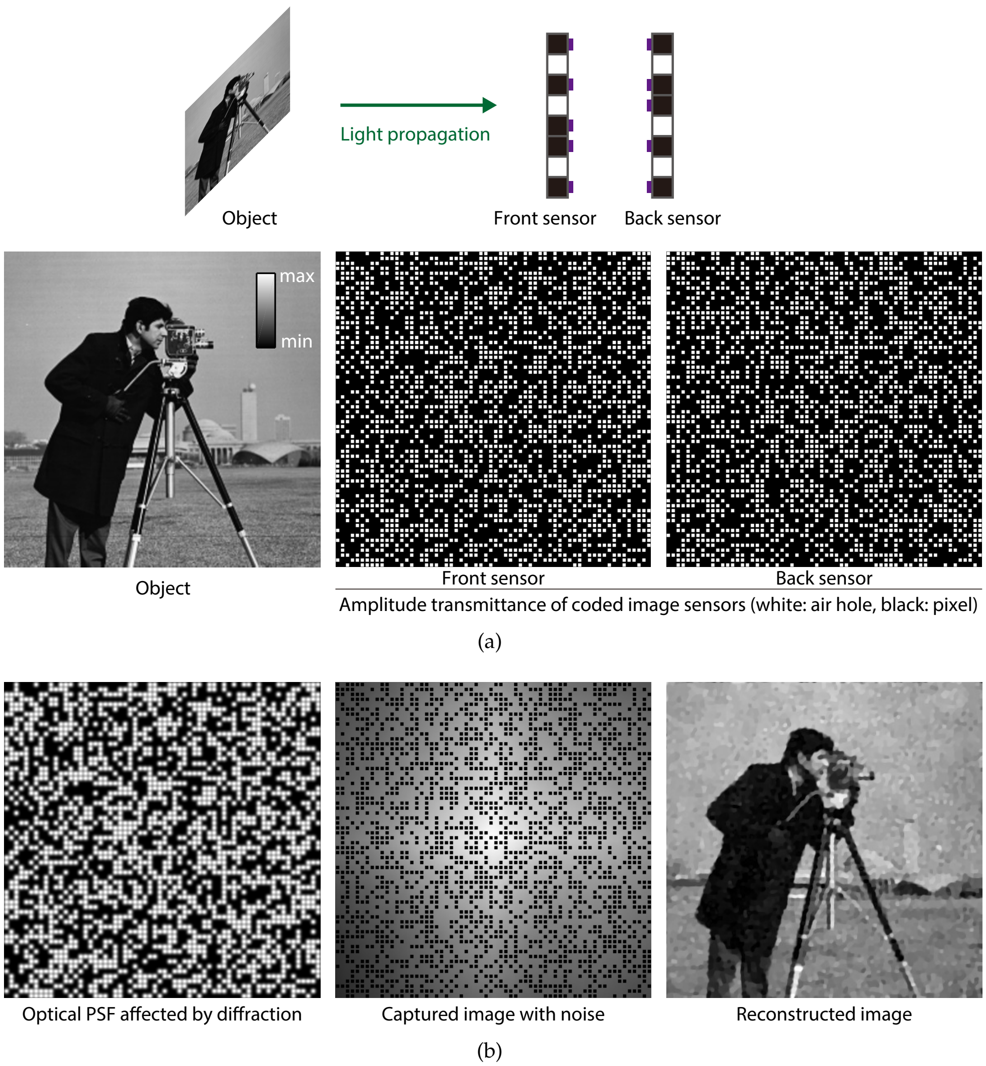


Figure 6. (a) Setup for imaging simulation including a subject and the amplitude transmittance of two CISs when the air-hole ratio was 50 %. (b) Results of the simulation including the optical point-spread function (PSF), a captured sparsely-sampled coded image with noise, and the reconstructed image.

The subject and the amplitude transmission of both the front and back sensors used for the simulation are shown in Fig. 6(a), together with an illustration of the assumed optical setup. In the figures of the image sensors, a white pixel physically means an air hole (drilled), whereas a black one indicates an opaque pixel (not drilled). In this simulation, we set the air-hole ratio to 50 % for both sensors. We designed the back sensor by rotating the front sensor by 90 degree, which enables differentiation of structures

Peer-reviewed version available at Sensors 2019, 19, 1329; doi:10.3390/s19061329

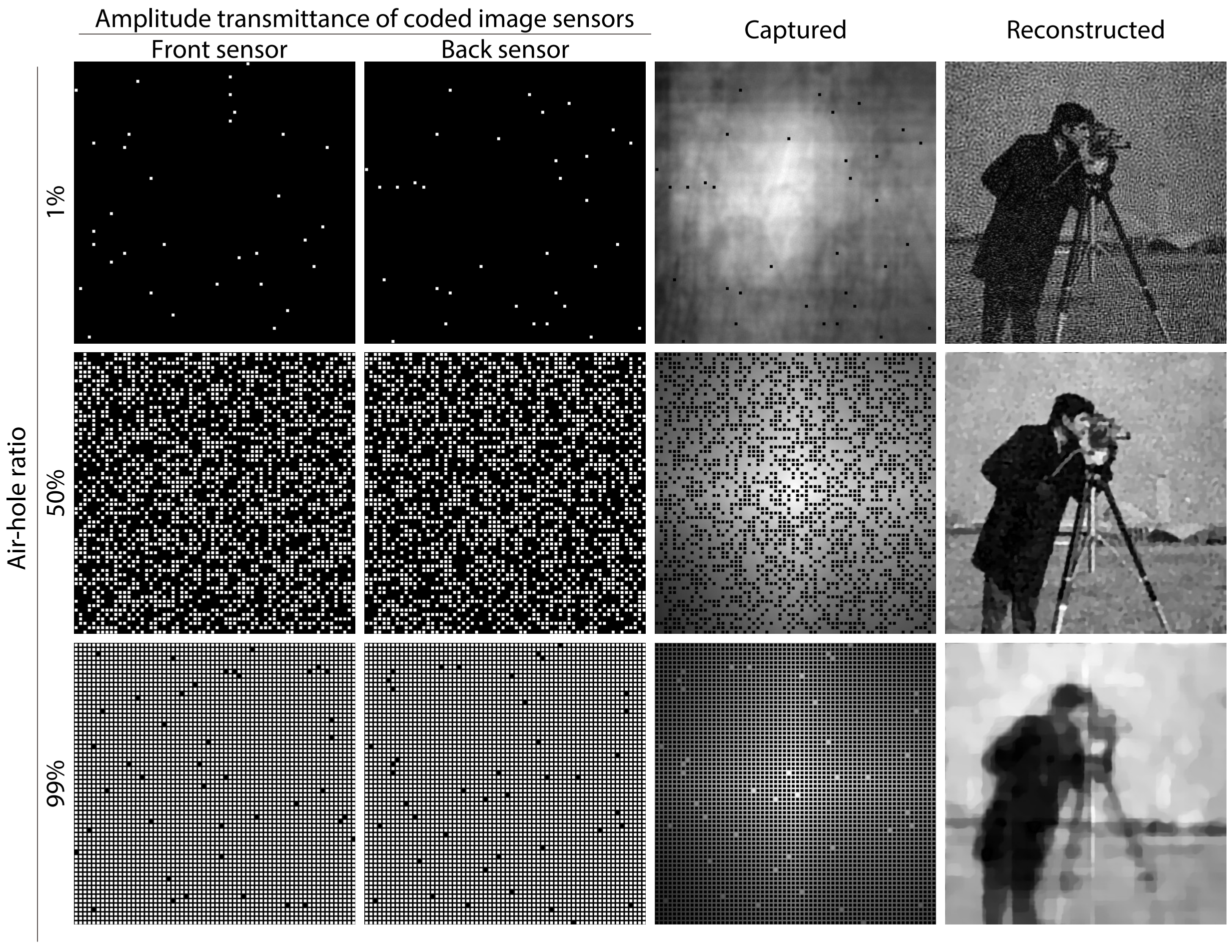
of two CISs manufactured with the same air-hole structure. As the subject, we used a standard image "Cameraman".

The simulation results are shown in Fig. 6(b). The optical PSF was a blurred image of the amplitude transmittance of the front sensor. To emulate photon-shot noise, the sparse-coded image was multiplied by Poisson noise, where the simulated mean signal-to-noise ratio (SNR) was 51.6 dB. As expected, the appearance of the captured image was not similar to the original object; nevertheless, the image information was encoded inside the image, though it was sparsified. The reconstructed image is shown in the right column of Fig. 6(b). For image decoding, τ in Eq. (5) was adjusted to 1×10^{-6} experimentally. The number of iterations of the algorithm was set to 10000 to ensure convergence of the cost function in optimization. As a result, we confirmed that an image of the object was successfully reconstructed. The peak SNR (PSNR) of the reconstructed image was 24.7 dB. The total processing time was 3.82 minutes when running Matlab R2017b on a computer with an Intel Xeon E5-2697 v4 CPU, an NVIDIA Geforce GTX 1080Ti, and 128 GB RAM.

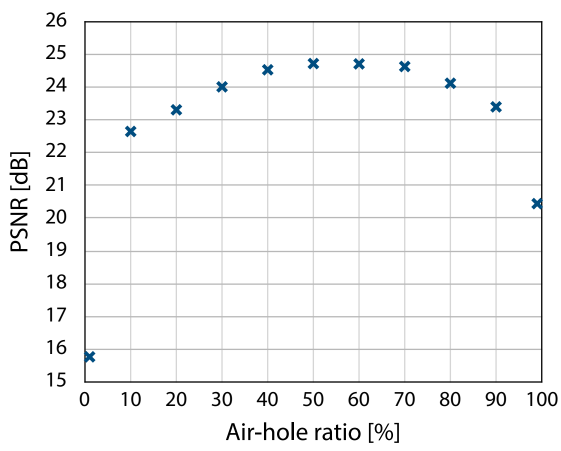
We also investigated the effect of the air-hole ratio in the image sensor. Here we assume that the front and back image sensors are implemented with the same air-hole ratio. In such case, there is a trade-off between the amount of transmissive light at the front sensor and the sampling count at the back sensor. Increasing the air-hole ratio results in a better SNR due to the higher detected photon count, however, it simultaneously results in reduced fidelity of the CS-based image reconstruction due to the sparser sensing. Therefore, the air-hole ratio should be designed optimally by considering the imaging situation, such as the environmental light intensity, the object's sparseness, and so on. Figure 7(a) shows examples of the reconstructed images obtained with different air-hole ratios. In the figure, the air-hole ratio was set to 1 %, 50 %, and 99 % from the left to the right. In this simulation, the change of the ratio directly affects the amount of photon-shot noise and the sampling count of the captured data. Note that the standard deviation of the photon-shot noise is equal to the square root of the detected photon count. Looking at the visual appearance in the figure, a lower ratio resulted in noisy-image reconstruction, whereas a higher ratio resulted in lower-fidelity reconstruction, as expected. Figure 7(b) shows the PSNRs of the reconstructed images obtained when changing the ratio from 1 % to 99 %. For example, the PSNR with ratios of 1 %, 50 %, and 99 % were 15.8 dB, 24.7 dB, and 20.4 dB, respectively. The results quantitatively indicated that an air-hole ratio of around 50 % resulted in the best final image quality under the example conditions used in the simulation.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 February 2019 doi:10.20944/preprints201902.0265.v1

Peer-reviewed version available at Sensors 2019, 19, 1329; doi:10.3390/s19061329



(a)



(b)

Figure 7. Simulation results obtained with different air-hole ratios of the CISs. (a) Used sensors, captured images, and reconstructed images when the air-hole ratio was 1 %, 50 %, and 99 %. (b) PSNRs of the reconstructed images, where the air-hole ratio ranged from 1 % to 99 %.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 February 2019 doi:10.20944/preprints201902.0265.v1

Peer-reviewed version available at Sensors 2019, 19, 1329; doi:10.3390/s19061329

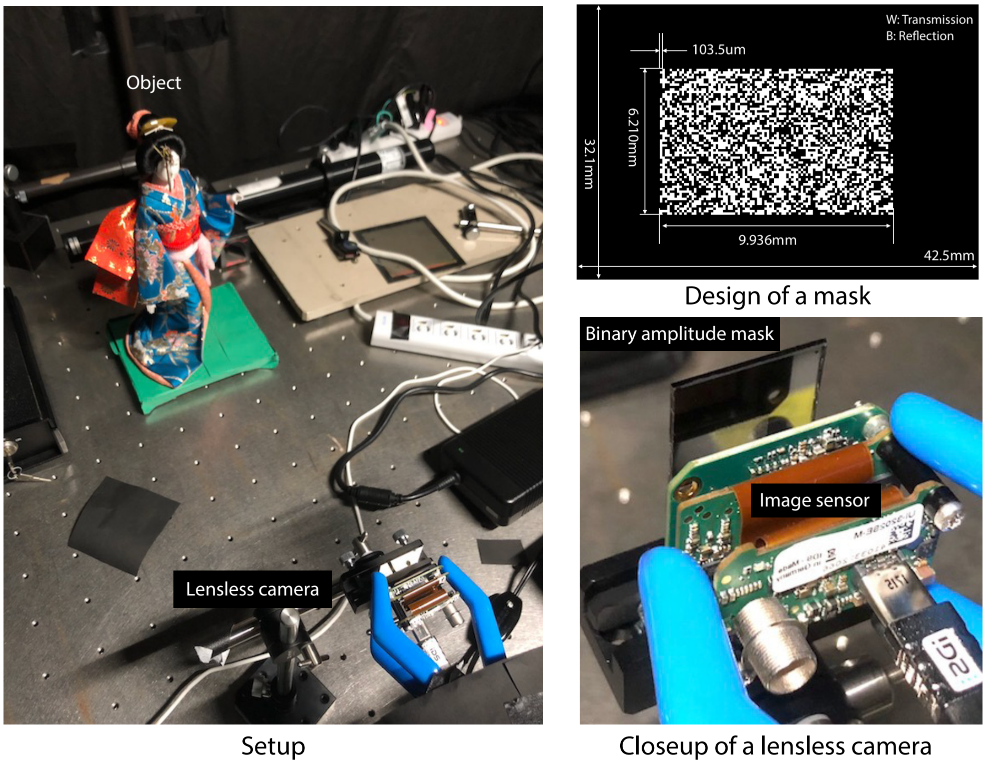


Figure 8. Experimental setup for capturing coded images using a mask-based lensless camera. The mask pattern indicates its amplitude transmission.

4. Experiments using a mask-based lensless camera and simulated air holes

We further investigated the validity of the proposed imaging method using data experimentally captured by a real mask-based lensless camera. Figure 8 shows the experimental setup. To implement a lensless camera, first, we designed and fabricated a binary amplitude mask with a randomly-coded structure. The mask was implemented by chromium deposition on a synthetic-silica plate. It was placed in front of a color CMOS image sensor (UI-3202SE-C by IDS) which contained 4104×3006 pixels with 3.45 μm pitch. Note that we resized the image to 512×512 pixels after the image capturing and demosaicing to accelerate the CS-based image decoding. The combination of the image sensor and the mask worked as the optical hardware of a lensless camera. As the subject, we placed a liquid-crystal display (LCD) or a doll in front of the lensless camera. The PSF of the lensless camera was calibrated experimentally by capturing the image of a point light source. When imaging, the space between the mask and the sensor was covered by light-shielding tape. In this experiment, random sparse sampling was computationally emulated. We will fabricate a real CIS in future work. For color-image acquisition, we performed image decoding for each color channel independently, where the decoding method for each channel was the same as that in the previous section. For simplicity, we assumed that the air holes were implemented pixel-by-pixel randomly.

Figure 9 shows the results of the experiment. As the subjects, we used a "Pepper" image displayed on an LCD and a Japanese doll as a diffuse object. After the capturing the raw coded image, image demosaicing was performed for generating a color coded image, and 50 % of the pixels were set to zero values to emulate the sensing by a CIS. The reconstructed images are shown in the right column of the figure. As a result of a visual assessment, the images of the subject were well reconstructed. The result validates the principle of the proposed imaging method in a real situation. The effects of using a real sparsified image sensor will be analyzed in future work.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 28 February 2019 doi:10.20944/preprints201902.0265.v1

Peer-reviewed version available at Sensors 2019, 19, 1329; doi:10.3390/s19061329

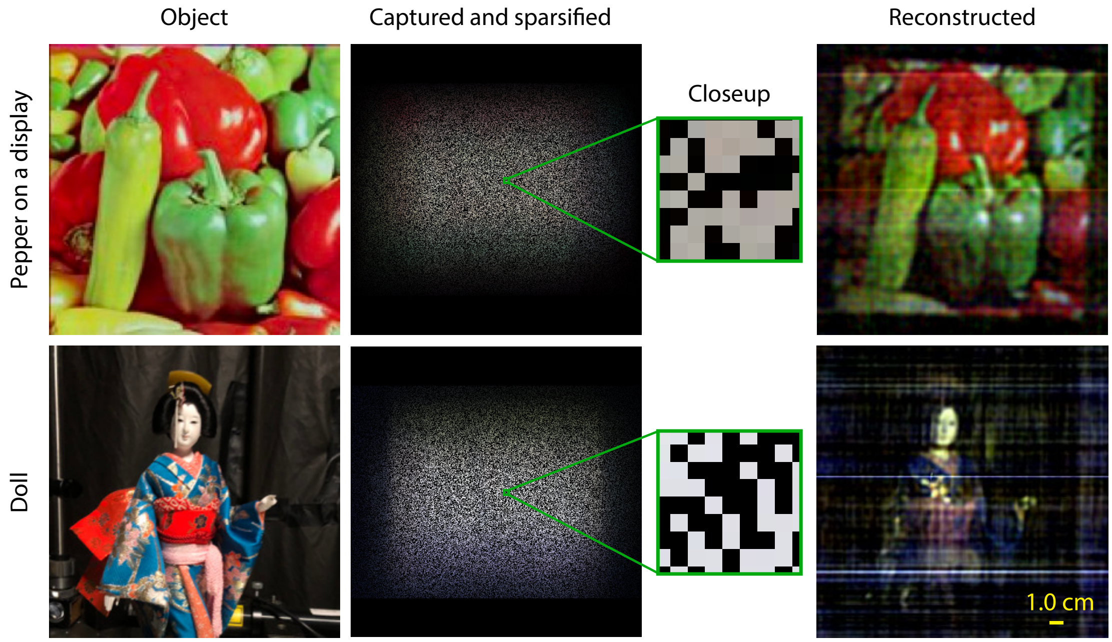


Figure 9. Experimental results using a real mask-based lensless camera and simulated air holes. The results includes subjects, captured and sparsified images, and reconstructed images. Scale bar indicates 1.0 cm.

In the reconstructed images, artifacts like a set of horizontal lines appeared. We assumed the artifacts were caused by the discontinuity of the captured coded image at the border of the pixel area in the image sensor. We also considered that the vignetting of the reconstructed image was caused by the design of the chief-ray angle (CRA) characteristics of each pixel in the consumer image sensor used, as mentioned also in [6]. This can be solved by controlling the CRA characteristics. In this simulation, we approximated the forward model of imaging as convolution with a space-invariant PSF in the decoding processing like [10]. However, in physical coded-image sensing, the PSF is not strictly space-invariant especially when the incident angle of light is large. We considered that the gap in such space-variance affected the degradation of the reconstructed image. In the experiments in this section, we repeated the same decoding processing three times independently, once for each color channel. However, aggressive utilization of the correlation of color-domain image information is a reasonable approach for improving the final image quality, as done in a computational compound-eye camera [19]. We will develop a decoding algorithm that exploits color-channel correlation in future work.

5. Discussion

In this paper, we proposed the novel architecture on the lensless camera that can capture the front and back scene at once. The merit of the proposed camera is the capability of super-FOV imaging with a thin and compact optical hardware constructed only by image sensors. The theoretical limit of the FOV for a single directional imaging is up to 180 degree, which is as same as that for a pinhole camera. However, the practical limit of the single FOV is restricted by the CRA characteristics of the detectors like experiments in this paper. This problem can be solved by appropriate control of the CRA characteristics of the detectors in manufacturing process. The FOV is also limited by the thickness of air holes mounted on the CIS. This limit can be solved by using more than two CISs with constructing regular-polyhedron-shape optical hardware, though discussion in this paper focused on the use of only two CISs for theoretical simplicity. For example, the use of six CISs with the construction of a cube-shape optical hardware is a promising approach to realize practical omnidirectional imaging. If the CMOS image sensor can be implemented spherically like Ref. [22], a single CIS is sufficient to realize the omnidirectional lensless imaging, which should be the smallest and the simplest realization.

To realize the proposed camera, the physical implementation of the CMOS CIS is necessary. In future works, such CMOS image sensor will be physically implemented, and the practical issues to use such a sensor will be addressed. The drilling of pixels on a CMOS image sensor itself was technically possible and demonstrated in past studies [20,21]. We regard that the same process can be used for implementing the proposed lensless camera. Afterward, the super-FOV imaging with the physically-implemented CISs will be demonstrated. The methods of optical encoding and computational decoding will also be improved by considering color-domain image correlation, aggressive design of the coding patterns, and so on.

The proposed camera can realize both compactness and wide-FOV, which can be applied to rich information sensing from severely-limited space such as endoscopy, inspection, robot vision, wearable sensing, and so on. The implication of the proposed novel camera ranges from the improvement of throughput of existing imagers to the pioneering new vision-technology applications.

6. Conclusions

In this paper, we proposed the novel architecture on the lensless camera that can capture the front and back scene at once. The merit of the proposed camera is the capability of super-FOV imaging by a thin and compact optical hardware constructed only by image sensors. To realize this, we exploited the CIS, which is the CMOS image sensor whose pixels are randomly drilled into air holes. In the proposed camera, CISs are placed facing with each other, where the front sensor works as a coded aperture and the back one works as a sparse image sampler of the coded optical image. The captured sparse coded image was computationally decoded into the subject image by the CS-based image-reconstruction algorithm.

We verified the proposed imaging system by simulations with numerical data and experiments with a real mask-based lensless camera and simulated air holes. The simulation and experimental results in this paper indicate the feasibility of the proposed lensless camera working with the sparse-sampling coded imaging. This indicates that the super-FOV imaging by a thin lensless camera, where both front and back scene can be captured at once, is possible using the proposed architecture of Fig. 1(b) if the CIS is physically implemented. The demonstration of the proposed super-FOV lensless camera using physically-implemented CMOS CIS will be a future work.

Author Contributions: conceptualization, T.N. and K.K.; methodology, T.N. and K.K.; software, T.N. and S.T.; validation, T.N. and S.T.; investigation, T.N.; writing–original draft preparation, T.N.; writing–review and editing, T.N.; visualization, T.N.; supervision, T.N., K.K. and M.Y.; project administration, T.N.; funding acquisition, T.N.

Funding: This research was funded by Precursory Research for Embryonic Science and Technology grant number JPM JPR1677 and Japan Society for the Promotion of Science grant number 18H03257.

Conflicts of Interest: The authors declare no conflict of interest.

Peer-reviewed version available at Sensors 2019, 19, 1329; doi:10.3390/s19061329

Abbreviations

The following abbreviations are used in this manuscript:

CIS	coded image sensor
CMOS	complementary metal–oxide–semiconductor
FOV	field-of-view
CS	compressive sensing
PSF	point-spread function
TV	total variation
FWHM	full-width half maximum
SNR	signal-to-noise ratio
PSNR	peak SNR
LCD	liquid-crystal display
CRA	chief-ray angle

References

1. Boominathan, V.; Adams, J.K.; Asif, M.S.; Avants, B.W.; Robinson, J.T.; Baraniuk, R.G.; Sankaranarayanan, A.C.; Veeraraghavan, A. Lensless Imaging: A computational renaissance. *IEEE Signal Proc. Mag.* **2016**, *33*, 23–35.
2. Brady, D.J. *Optical Imaging and Spectroscopy*; Wiley, 2009.
3. Stork, D.G.; Gill, P.R. Optical, Mathematical, and Computational Foundations of Lensless Ultra-Miniature Diffractive Imagers and Sensors. *Int. J. Adv. Syst. Meas.* **2014**, *7*, 201–208.
4. Horisaki, R.; Ogura, Y.; Aino, M.; Tanida, J. Single-shot phase imaging with a coded aperture. *Opt. Lett.* **2014**, *39*, 6466–6469.
5. Asif, M.S.; Ayremlou, A.; Sankaranarayanan, A.C.; Veeraraghavan, A.; Baraniuk, R.G. FlatCam: Replacing Lenses with Masks and Computation. Proceedings of the 2015 IEEE International Conference on Computer Vision Workshop, Santiago, Chile, 11–18 December 2015.
6. Shimano, T.; Nakamura, Y.; Tajima, K.; Sao, M.; Hoshizawa, T. Lensless light-field imaging with Fresnel zone aperture: quasi-coherent coding. *Appl. Opt.* **2018**, *57*, 2841–2850.
7. Potuluri, P.; Xu, M.; Brady, D.J. Imaging with random 3D reference structures. *Opt. Express* **2003**, *11*, 2134–2141.
8. Nakamura, T.; Horisaki, R.; Tanida, J. Compact wide-field-of-view imager with a designed disordered medium. *Opt. Rev.* **2015**, *22*, 19–24.
9. Kim, G.; Isaacson, K.; Palmer, R.; Menon, R. Lensless Photography with only an image sensor. *Appl. Opt.* **2017**, *56*, 6450–6456.
10. Antipa, N.; Kuo, G.; Reinhard, H.; Mildenhall, B.; Bostan, E.; Ng, R.; Waller, L. DiffuserCam : lensless single-exposure 3D imaging. *Optica* **2018**, *5*, 1–9.
11. Donoho, D.L. Compressed Sensing. *IEEE Trans. Inf. Theory* **2006**, *52*, 1289–1306.
12. Candès, E.; Wakin, M. An Introduction To Compressive Sampling. *IEEE Signal Proc. Mag.* **2008**, *25*, 21–30.
13. Stern, A. *Optical Compressive Imaging*; CRC Press, 2016.
14. Nakamura, T.; Kagawa, K.; Torashima, S.; Yamaguchi, M. Lensless imaging by coded image sensors. Proceedings of the 4th International Workshop on Image Sensors and Imaging Systems, Tokyo, Japan, 28–29 November 2018.
15. Rudin, L.I.; Osher, S.; Fatemi, E. Nonlinear total variation based noise removal algorithms. *Physica D* **1992**, *60*, 259–268.
16. Candès, E.; Romberg, J. Sparsity and incoherence in compressive sampling. *Inverse Probl.* **2007**, *23*, 969–985.
17. Goodman, J.W. *Introduction to Fourier Optics*; McGraw-Hill, 1996.
18. Bioucas-Dias, J.M.; Figueiredo, M.A.T. A New TwIST : Two-Step Iterative Shrinkage / Thresholding Algorithms for Image Restoration. *IEEE Trans. Image Process.* **2007**, *16*, 2992–3004.
19. Venkataraman, K.; Lelescu, D.; Duparré, J.; McMahon, A.; Molina, G.; Chatterjee, P.; Mullis, R.; Nayar, S.K. PiCam: An Ultra-Thin High Performance Monolithic Camera Array. *ACM Trans. Graph.* **2013**, *32*, 166.

20. Ng, D.C.; Tamura, H.; Mizuno, T.; Tokuda, T.; Nunoshita, M.; Ishikawa, Y.; Shiosaka, S.; Ohta, J. An implantable and fully integrated complementary metal-oxide semiconductor device for in vivo neural imaging and electrical interfacing with the mouse hippocampus. *Sensor Actuat. A-Phys.* **2008**, *145-146*, 176–186.
21. Ohta, J.; Tokuda, T.; Sasagawa, K.; Noda, T. Implantable CMOS Biomedical Devices. *Sensors* **2009**, *9*, 9073–9093.
22. Zhang, K.; Jung, Y.H.; Mikael, S.; Seo, J.H.; Kim, M.; Mi, H.; Zhou, H.; Xia, Z.; Zhou, W.; Gong, S.; Ma, Z. Origami silicon optoelectronics for hemispherical electronic eye systems. *Nat. Commun.* **2017**, *8*, 1782.