

ARTICLE

Open Access

# Single-shot lensless imaging with fresnel zone aperture and incoherent illumination

Jiachen Wu<sup>1</sup>, Hua Zhang<sup>1</sup>, Wenhui Zhang<sup>1</sup>, Guofan Jin<sup>1</sup>, Liangcai Cao<sup>1</sup> <sup>1</sup> and George Barbastathis<sup>2</sup>

## Abstract

Lensless imaging eliminates the need for geometric isomorphism between a scene and an image while allowing the construction of compact, lightweight imaging systems. However, a challenging inverse problem remains due to the low reconstructed signal-to-noise ratio. Current implementations require multiple masks or multiple shots to denoise the reconstruction. We propose single-shot lensless imaging with a Fresnel zone aperture and incoherent illumination. By using the Fresnel zone aperture to encode the incoherent rays in wavefront-like form, the captured pattern has the same form as the inline hologram. Since conventional backpropagation reconstruction is troubled by the twin-image problem, we show that the compressive sensing algorithm is effective in removing this twin-image artifact due to the sparsity in natural scenes. The reconstruction with a significantly improved signal-to-noise ratio from a single-shot image promotes a camera architecture that is flat and reliable in its structure and free of the need for strict calibration.

## Introduction

Traditional optical imaging architectures follow a point-to-point imaging model by using a set of lenses. This way to collect and converge light limits the freedom of imaging parameters. By introducing computing capability into the imaging system, breakthrough systems have been built, yielding improvements in the image dimensions<sup>1–3</sup>, image size<sup>4,5</sup>, and even imaging mechanisms<sup>6–8</sup>. Furthermore, the burden of imaging can be transferred from bulky and expensive hardware to computation, which enables new architectures for low-cost cameras.

Recently, lensless imaging has become attractive due to its thin and easy-to-build form. In the past few years, various lensless imaging techniques have been proposed with coherent systems, such as an on-chip microscope<sup>9,10</sup>, coherent diffractive imaging<sup>11,12</sup>, and a series of learning-based methods<sup>13–15</sup>. Due to the requirement of coherent illumination, however, applications of such systems are

limited. The pinhole camera offers a simple architecture for incoherent light imaging, which has an infinite depth of field and no aberration but suffers from the severe limitation of low light throughput. To make up for this drawback, coded aperture imaging replaces the pinhole with a mask. Each light source in the scene casts a unique shadow of the mask onto the sensor, encoding the intensity and the location of the light source. Masks, which include a uniform redundant array<sup>16</sup> or modified URA<sup>17</sup>, have traditionally been used in X-ray and  $\gamma$ -ray imaging. Several lensless cameras at visible wavelengths have been proposed using diffractive gratings that are insensitive to wavelength<sup>18,19</sup>, separable masks<sup>20,21</sup> and diffusers<sup>22</sup>. These methods depend on large-scale matrix inversion, which requires strict calibration and heavy computational resources.

Unlike the pure mathematical optimization methods mentioned above, we adopt the Fresnel zone plate as the mask to improve the condition of the imaging problem at the physical level. Soon after Gabor's invention of holography<sup>23</sup>, Rogers noted that the pattern of the Fresnel zone plate coincides with a point source hologram<sup>24</sup>. Inspired by this, Mertz and Young proposed zone plate coded imaging<sup>25</sup> and extended the concept

Correspondence: Liangcai Cao ([clc@tsinghua.edu.cn](mailto:clc@tsinghua.edu.cn)) or George Barbastathis ([gbarb@mit.edu](mailto:gbarb@mit.edu))

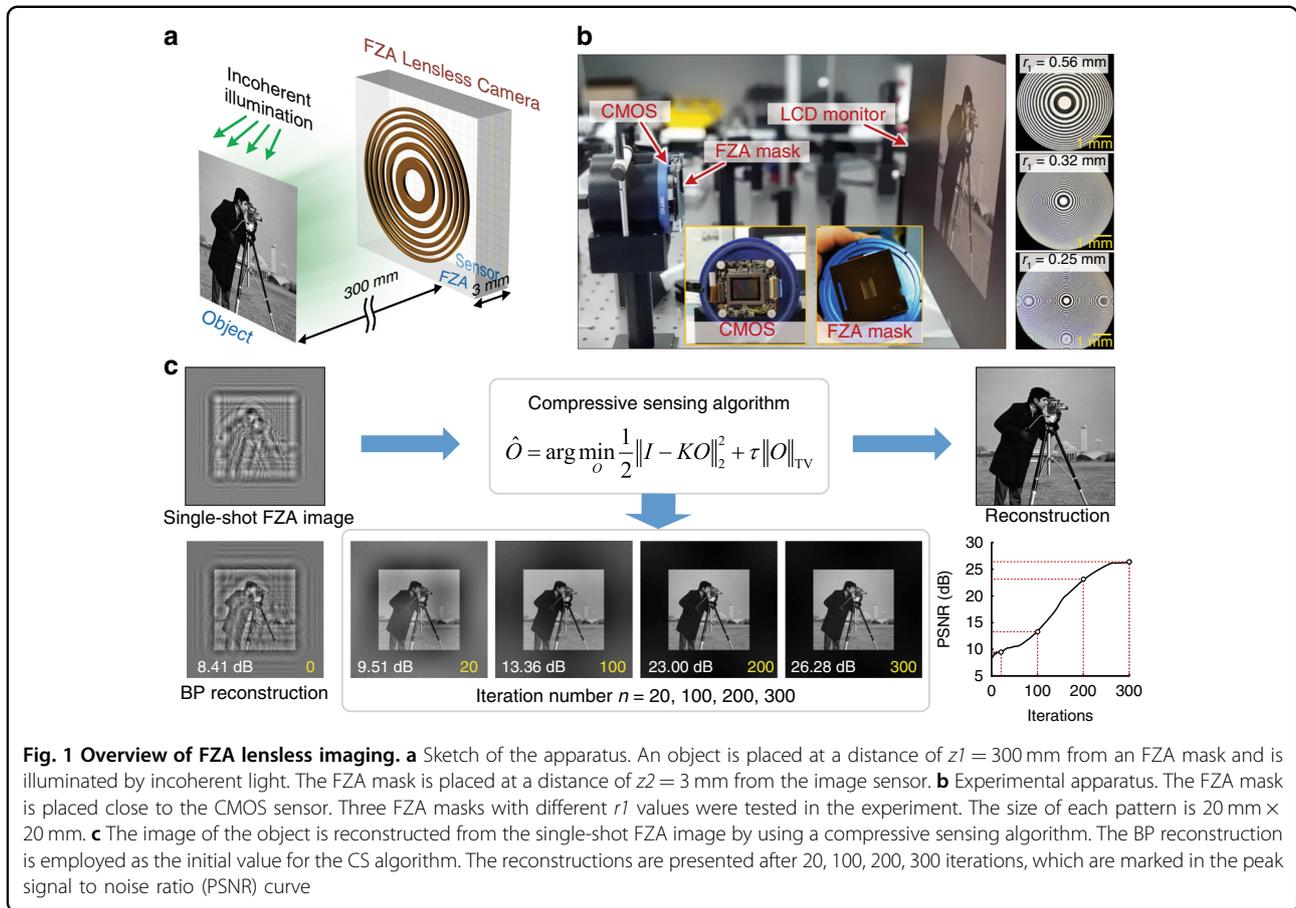
<sup>1</sup>State Key Laboratory of Precision Measurement Technology and Instruments, Department of Precision Instruments, Tsinghua University, 100084 Beijing, China

<sup>2</sup>Department of Mechanical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

© The Author(s) 2020



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.



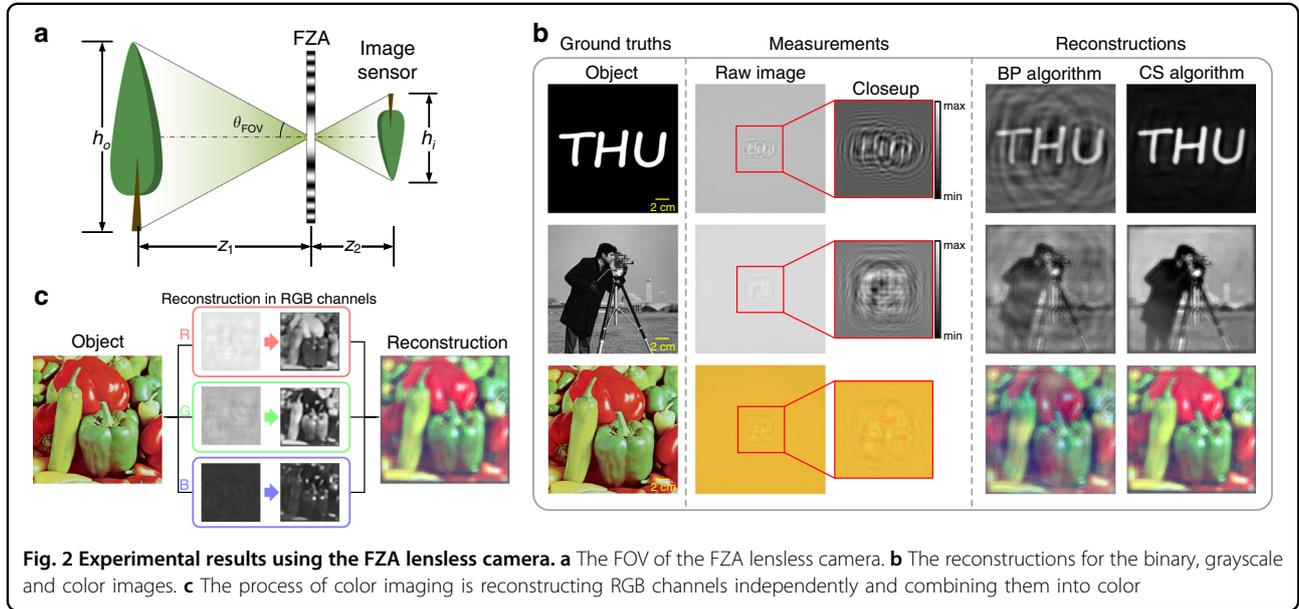
and application of holography to the field of incoherent light. Recent works have demonstrated the imaging capability of the Fresnel zone plate, also called the Fresnel zone aperture (FZA), at visible wavelengths<sup>26–29</sup>. This prototype camera needs at least four FZAs with different phases to extract the signal from captured images. Time division and spatial division are two alternatives to implement the change of masks. The former increases the complexity of the system by using a spatial light modulator. The latter reduces the resolution because the size of the reconstruction image is only a quarter of the sensor size. Moreover, the crosstalk of light and sensor noise would further degrade the quality of the reconstructed image.

In mask-based imaging systems, image reconstruction is susceptible to noise, so a robust reconstruction algorithm is critical. Compressive sensing (CS) is a powerful signal reconstruction framework and provides complete theoretical support for image reconstruction<sup>30</sup>. The CS algorithm has been widely applied in optical imaging systems, such as single pixel cameras<sup>31–33</sup> and compressive holography<sup>34–38</sup>. These previous works pave the way for inverse problem solving in optical imaging and inspire us in terms of image reconstruction.

In this work, we propose a single-shot noise suppression lensless imaging method with FZA and incoherent illumination. The structure of the proposed imaging system is simple: an FZA mask is placed close to a sensor (Fig. 1a) without precise alignment. Each point source in the scene casts an FZA shadow on the sensor plane, which is similar to the point source hologram, so that the FZA acts in some sense as a Fresnel hologram encoder. Backpropagation (BP) reconstruction is feasible but disturbed by twin image. Here, the CS algorithm that has been proven to be efficient for Fresnel holography is applied (Fig. 1c). By enforcing a sparsity constraint in the gradient domain, the reconstruction eliminates the twin-image effect and suppresses the other noise. Since the single-shot method only needs a single FZA, all the pixels can be saved for a full-size reconstruction image, while in the multiple-shot method, only a quarter size can be realized. The proposed method can greatly support high-quality imaging without the need for strict calibration for the FZA lensless camera.

### Results

A lensless camera using a QHY163M CMOS image sensor is designed and arranged as shown in Fig. 1b. The size of the image sensor is  $17.7 \text{ mm} \times 13.4 \text{ mm}$ , and the number of



pixels is  $4656 \times 3522$ . The pixel pitch is  $3.8 \mu\text{m}$ . The captured image is cropped to  $2048 \times 2048$  pixels for subsequent processing. An FZA mask is placed close to the CMOS sensor. The transmission function of the FZA mask is

$$T(r) = \frac{1}{2} + \frac{1}{2} \text{sgn} \left[ \cos \left( \frac{\pi r^2}{r_1^2} \right) \right] \quad (1)$$

where  $r_1$  denotes the radius of the innermost zone and  $r$  is the radial distance from the center of the aperture. In addition, “sgn” is the signum function that is  $+1$  for a positive number or  $-1$  for a negative number. We use three FZAs with different  $r_1$  to test the imaging resolution. The thickness of the photomask is  $2 \text{ mm}$ , and the thickness of the protective glass is  $1 \text{ mm}$ . Thus, the distance between the mask and the sensor is  $3 \text{ mm}$ . An LCD monitor with  $1920 \times 1080$  resolution is placed  $\sim 300 \text{ mm}$  from the FZA. The test images are displayed on the screen, and the sensor records the rays from the screen that are modulated by the mask. For each test image, we adjusted the exposure time according to the scene intensity to avoid overexposure or underexposure.

The FZA imaging follows the pinhole imaging model. The position of each reconstructed image point on the sensor plane is at the intersection of the chief ray passing through the center of the FZA and the sensor plane (Fig. 2a). The magnification of the system is derived from the geometrical relationship:

$$M = \frac{h_i}{h_o} = \frac{z_2}{z_1} \quad (2)$$

where  $h_o$  is the object height and  $h_i$  is the image height. The field of view (FOV) is expressed as

$$\theta_{\text{FOV}} = \arctan \left( \frac{h_i}{2z_2} \right) \quad (3)$$

The binary, grayscale and color images are tested in our experiments. The original images, the corresponding measured images, and the reconstruction results are shown in Fig. 2b. The original image is  $20 \text{ cm} \times 20 \text{ cm}$  when displayed on screen. The measured images are monochrome images with 16-bit depth. For binary and grayscale images, the closeups of the raw images normalize the grayscale of the measured images. The BP and CS algorithms are used to reconstruct images from measured images. According to the magnification, the central  $526 \times 526$  portions of each image are cropped to be shown in Fig. 2b.

The FZA lensless camera has the capacity for color imaging as well. Because the imaging model is based on geometrical optics, the reconstruction is not influenced by the wavelength of the object. By using an RGB sensor, the intensity of different wavelength bands can be recorded by the RGB channels independently and the reconstruction algorithm applied to the three channels. Then, the reconstructed images of the three channels are combined into the final color image (Fig. 2c).

### Discussion Resolution

In practice, the sensor and the FZA mask both have finite sizes, which means that the reconstruction is band-limited.

If the sensor can completely record the FZA shadow and the pixel pitch is small enough to satisfy the sampling theorem, the frequency range is limited by the number of recorded zones, which in turn depends on the radius of the FZA. The larger the radius, the finer the zones that are included. Thus, the resolution of the reconstructed image can be improved by increasing the aperture radius  $R$ . On the other hand, the FZA constant  $r_1$  also determines the number of zones within a fixed aperture radius  $R$ . Improving the resolution of the reconstructed image can also be realized by shrinking the FZA constant  $r_1$ .

Quantitative analysis is carried out by means of the coherent impulse response (CIR). As shown in section “Resolution”, the CIR of the imaging system is calculated as

$$I_{\text{CIR}}(r_o) = \exp\left(\frac{i\pi}{r_1^2} r_o^2\right) \frac{R}{r_0} J_1(2\pi r_o R / r_1^2) \quad (4)$$

where  $R$  is the radius of FZA and  $J_1(\cdot)$  is the first-order Bessel function of the first kind. Because the intensity of the object is real-valued, only the real part of the CIR should be considered. According to the Rayleigh criterion, the minimum distance between resolvable points is defined as the distance from the center to the first zero of the CIR. The first zero of the real part of the exponential term is  $0.707r_1$ , whereas the first zero of the order-one Bessel function is  $0.61(r_1/R)r_1$ . Since  $r_1 \ll R$ , the resolution is determined by

$$r_c = 0.61 \frac{r_1^2}{R} \quad (5)$$

Assuming the FZA contains  $n$  zones, namely,  $R = \sqrt{n}r_1$ , then the width of the outermost zone is  $\Delta r = (\sqrt{n} - \sqrt{n-1})r_1 \approx r_1/(2\sqrt{n})$ . In terms of  $\Delta r$ , the resolution is expressed as

$$r_c = 1.22\Delta r \quad (6)$$

Equation (6) reveals the simple yet useful result that the resolution of an FZA imaging system is approximately equal to the width of the outermost zone. Taking the system magnification  $M$  into account, the resolution at the object plane is  $r'_c = (1/M)r_c$ .

Figure 3 shows the CIRs and the corresponding reconstructed images with different values of  $r_1$ . The aperture radius  $R = 5.12$  mm, and the values of  $r_1$  are 0.8, 0.5, and 0.3 mm. The corresponding values of  $r_c$  and  $\Delta r$  are labeled in Fig. 3, which agrees with Eq. (6). The smaller FZA constant  $r_1$  yields higher-quality reconstructions, which verifies the above discussion. Note that the resolution improvement is limited because when the

FZA constant decreases, diffraction becomes noticeable, which obfuscates the reconstruction.

A negative USAF 1951 test chart is displayed on the monitor to experimentally test the resolution. The displayed size has a magnification three times larger than the standard size. Three FZAs with the same aperture radius  $R$  and different FZA constants  $r_1$  are used for comparison. The FZA constants  $r_1$  are 0.56, 0.32, and 0.25 mm. The aperture size is approximately equal to the sum of the whole image size and object imaging size, and the aperture radius is 4.56 mm. The results are shown in Fig. 4. For  $r_1 = 0.56$  mm, group-2/element 5 can barely be resolved (Fig. 4b), which represents that the minimum resolved distance is 3.78 mm at the object plane (or 0.038 mm at the image plane, after multiplying by the magnification 0.01). For  $r_1 = 0.32$  mm, group 0/element 1 can barely be resolved (Fig. 4c), and the corresponding minimum resolved distance is 1.5 mm in the object plane and 0.015 mm in the image plane. The experimental values are close to the theoretical values, which are 0.042 mm and 0.014 mm calculated by Eq. (5), respectively. When  $r_1$  reduces to 0.25 mm, the resolution does not improve significantly (Fig. 4d); that is, the diffraction effect limits the resolution.

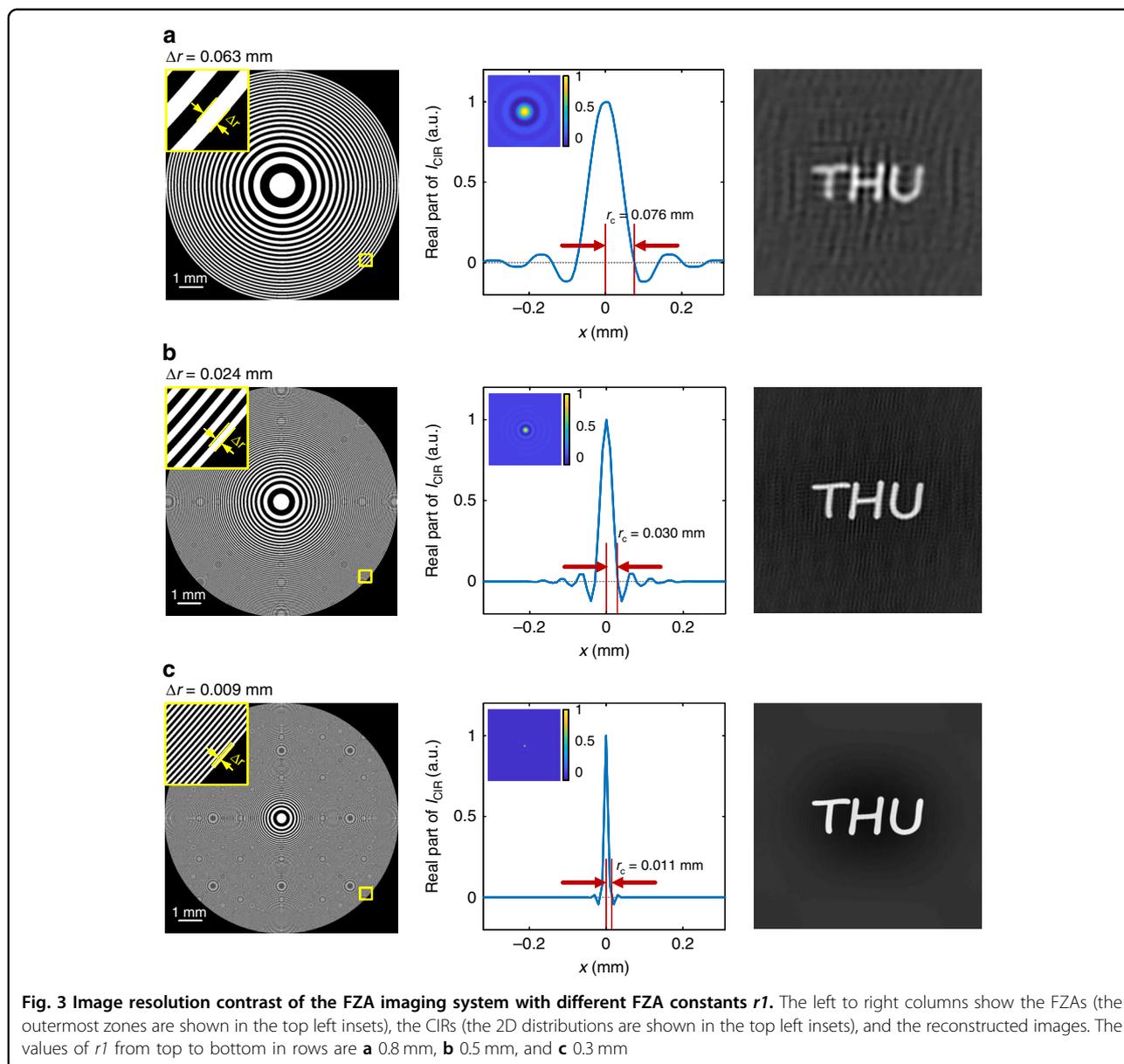
### Noise and error analysis

Except for the twin-image noise, the measurement can be corrupted by other noises and errors originating mainly from three sources: sensor, mask, and diffraction. The sensor noise consists of quantization noise and dark current noise. The effect of this noise on the reconstruction is relatively moderate and can be reduced by improving bit depth and increasing exposure time.

The mask error is caused by binarization of transmittance. The binary FZA, as a substitute of the GZP, consists of only transparent and opaque zones. However, in comparison to the GZP, which has only one pair of conjugate foci at  $f = \pm r_1^2/\lambda$ , the FZA has multiple foci with corresponding focal lengths  $f = \pm r_1^2/n\lambda$ ,  $n = 1, 3, 5, \dots$ . The property can be explained by using Fourier expansion. The Fourier series representation of this binary function is

$$T(r) = \frac{1}{2} + \frac{2}{\pi} \sum_{n=1}^{\infty} \frac{1}{n} \sin\left(\frac{n\pi}{2}\right) \cos\left(\frac{\pi n r^2}{r_1^2}\right) \quad (7)$$

The derivation can be seen in the supplementary information (Section S1). Equation (7) states that the intensity transmittance function of the FZA is a linear superposition of a series of GZPs with different focal lengths. The higher-order focal lengths cause a series of defocused images at the first-order focal plane, which degrades the image quality. This issue can be remedied by using a binary Gabor zone plate (GZP) that has only one pair of conjugate foci<sup>39,40</sup>.



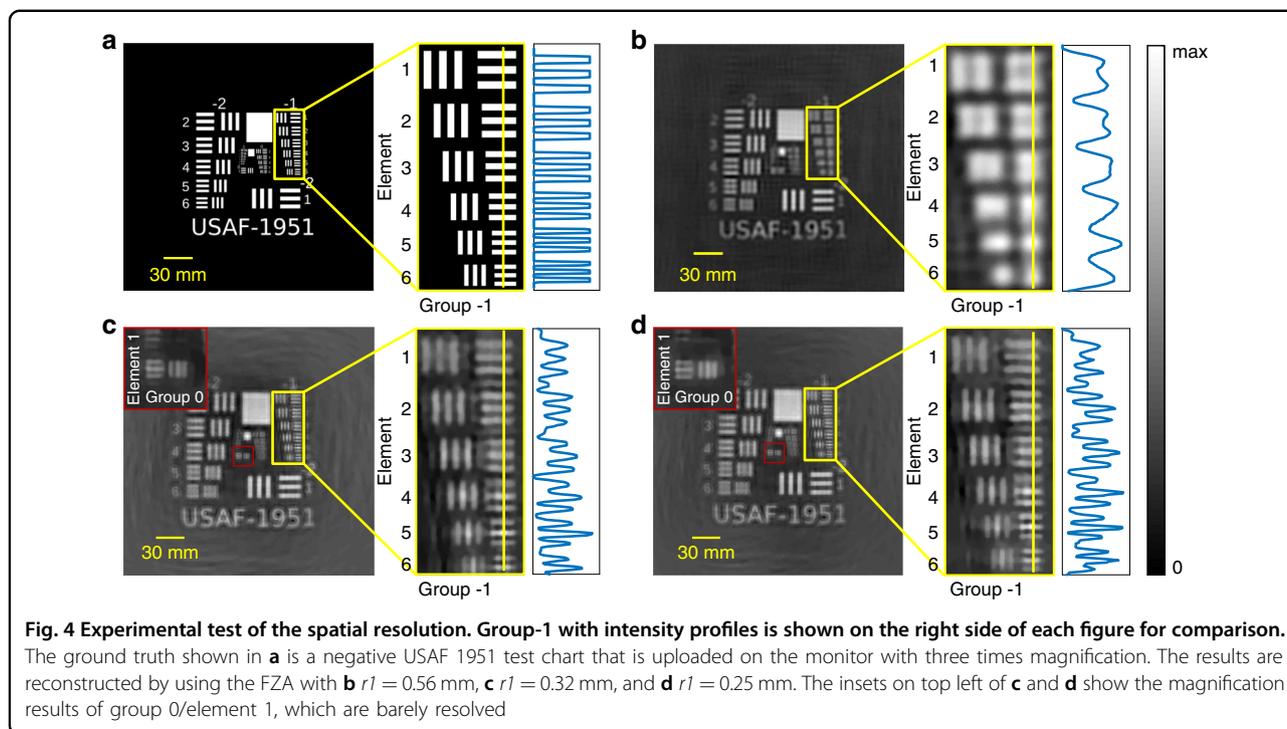
Increasing the regularization coefficient appropriately can also effectively suppress these defocused noises.

With a decreasing zone width, diffraction effects become noticeable. The diffracted pattern is no longer an ideal FZA pattern. To observe the diffraction phenomena, a white collimating LED light is adopted to illuminate the FZA mask. The FZA pattern with  $r_1 = 0.32$  mm and the corresponding diffracted pattern are shown in Fig. 5a. The cross section of the radial intensity is shown in Fig. 5b. The contrast drops to the lowest value around a radius of 500 pixels away from the center due to the diffraction propagation law of objects with periodic transmittance. Since the contrast of the diffraction fringe of grating

changes periodically along with propagation distance, the contrast is minimized when the propagation distance is an odd number of half Talbot lengths<sup>41</sup>; that is,

$$z^{(\min)} = l + \frac{1}{2}z_T = l + \frac{p^2}{2\lambda}, \quad l = 0, 1, 2, \dots \quad (8)$$

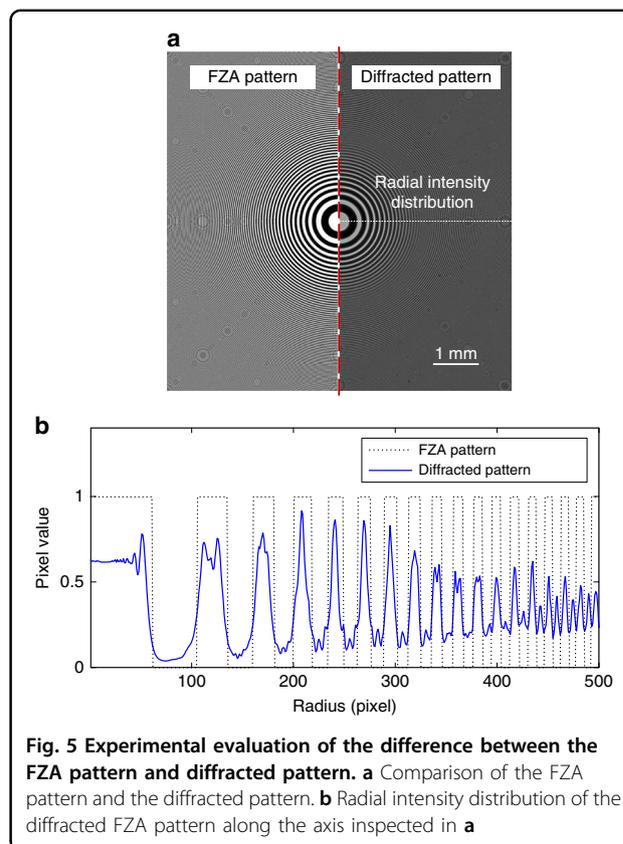
where  $p$  is the grating period. The distance is fixed, but the period  $p = r_1^2/r$  varies with the radius. Thus, the contrast also varies with the radius. This decrease in contrast results in model error and limits the reconstruction resolution. The mask can be improved by adopting a well-designed diffraction optical element to obtain the desired diffraction pattern in future work.

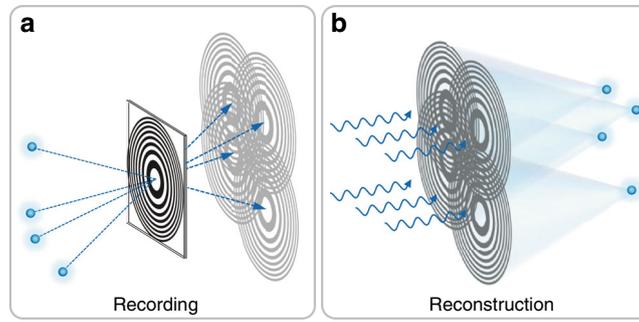


**Outlook**

We have proposed an FZA lensless imaging method under incoherent illumination using computational reconstruction. CS theory provides a guarantee for accurate signal reconstruction. The proposed camera consists of an image sensor and an FZA mask and thus is thin and compact. The incoherent rays pass through the FZA and cast object-dependent shadows on the sensor plane to compose the raw image. This mechanism is somewhat analogous to that of inline holography. Unlike conventional BP, which produces twin images, the compressive sensing algorithm with total variation (TV) regularization that we adopt here eliminates the twin-image effect and reconstructs the image with reduced noise.

The imaging resolution of this mask-based camera still has much room for further improvements by modifying the mask so that the diffracted pattern coincides with the expected pattern. The prospect of FZA imaging using such a relatively thin setup can lead to all kinds of applications. One is a portable camera to image various objects from a person in a room to the landscape outdoors. This approach can fulfill a range of recognition tasks, such as object detection, character recognition, and face recognition. Furthermore, the proposed camera can be monolithically fabricated by depositing the FZP pattern on the cover glass of the sensor so that it can be readily integrated with portable devices or any flat surface. It is suggested that this





**Fig. 6 Recording and reconstruction of FZA imaging.** **a** Each point source in the scene casts an FZA shadow on the sensor plane. These shadows are superimposed and form the pattern with the same form as the inline hologram. **b** Applying coherent propagation for the recorded pattern can reconstruct the point sources

ultrathin and low-cost camera has great potential with the development of computational imaging.

**Materials and methods**  
**Imaging model**

An ideal zone plate is called a GZP. Its amplitude transmission function is

$$T(r) = \frac{1}{2} + \frac{1}{2} \cos\left(\frac{\pi r^2}{r_1^2}\right) \tag{9}$$

However, such a zone plate is difficult to manufacture because of its sinusoidal variation transmittance. The FZA with binary transmission is a more practical alternative mask. For the sake of derivation, we substitute the transmission function of the GZP for the FZA in subsequent content. The object is placed at a distance of  $z_1$  from the FZA and illuminated by an incoherent light source. The FZA is placed in front of an image sensor at distance  $z_2$ . The object surface diffuses the light and can be considered a superposition of point sources. Each point source casts an FZA shadow on the sensor plane (Fig. 6a). The shadow center is at the intersection of the chief ray and the sensor plane. The size of the shadow is expanded from the FZA by the magnification factor  $(1 + z_2/z_1)$ . Then, the FZA constant of the shadow would be  $r_1' = (1 + z_2/z_1)r_1$ . Thus, the image formed on the sensor is a superposition of shifted and scaled versions of FZA shadows. When  $z_1 \gg z_2$ , the FZA constant  $r_1'$  is almost equal to  $r_1$ . In this case, the imaging process can be formulated as

$$I(\mathbf{r}) = \frac{1}{2} \sum_k^N I_k \left[ 1 + \cos\left(\frac{\pi}{r_1^2} |\mathbf{r} - \mathbf{r}_k|^2\right) \right] \tag{10}$$

where  $I(\mathbf{r})$  is the intensity distribution on the sensor plane;  $I_k$  is the intensity of the  $k$ th point source; the vector  $\mathbf{r}$  is the arbitrary position vector, and  $\mathbf{r}_k$  is the translation vector of the  $k$ th shadow in the sensor plane. Each FZA shadow can be considered a point source hologram that encodes the

intensity and the location of the point source. All these elementary holograms synthesize the final measurement. Then, the reconstruction can be performed by coherent propagation (Fig. 6b). Both optical and computational methods are available for the reconstruction.

In the Fresnel approximation, the reconstructed wavefront may be written as

$$O_R(\mathbf{r}_o) = \frac{\exp(i2\pi d/\lambda)}{i\lambda d} \iint I(\mathbf{r}) \exp\left[\frac{i\pi}{\lambda d} |\mathbf{r} - \mathbf{r}_o|^2\right] dS \tag{11}$$

where  $\lambda$  and  $d$  represent the wavelength of the reconstructed wavefront and the reconstructed distance, respectively. To ensure correct reconstruction, the two parameters should satisfy  $r_1^2 = \lambda d$ . The vector  $\mathbf{r}_o$  is the position vector in the reconstructed plane. Here,  $dS$  is the area element in the measured image. For the ideal case, the integral area is infinite. Expanding the cosine term in Eq. (10) into an exponential form and substituting it into Eq. (11) while ignoring the constant coefficient, Eq. (11) becomes

$$\begin{aligned} O_R(\mathbf{r}_o) &= \frac{1}{2} \iint \exp\left(\frac{i\pi}{r_1^2} |\mathbf{r} - \mathbf{r}_o|^2\right) dS \cdot \sum_k^N I_k \\ &\quad + \frac{1}{4} \sum_k^N I_k \iint \exp\left[\frac{i\pi}{r_1^2} (|\mathbf{r} - \mathbf{r}_o|^2 - |\mathbf{r} - \mathbf{r}_k|^2)\right] dS \\ &\quad + \frac{1}{4} \sum_k^N I_k \iint \exp\left[\frac{i\pi}{r_1^2} (|\mathbf{r} - \mathbf{r}_o|^2 + |\mathbf{r} - \mathbf{r}_k|^2)\right] dS \\ &= \frac{i\pi^2}{2} \sum_k^N I_k + \frac{r_1^4}{4} \sum_k^N I_k \delta(\mathbf{r}_o - \mathbf{r}_k) \\ &\quad + \frac{i\pi^2}{8} \sum_k^N I_k \exp\left(\frac{i\pi}{2r_1^2} |\mathbf{r}_o - \mathbf{r}_k|^2\right) \end{aligned} \tag{12}$$

The first term is a constant term that is proportional to the total intensity of the object. The second term is a set of points appearing at the same locations as the

geometrical imaging point, and the intensities are proportional to the intensities of the original light sources. These points reproduce the image of the original object. The third term is the superposition of spherical waves propagating from distance  $2d$ . It can be considered an out-of-focus image, which is the so-called twin image. The inherent twin image obscures the reconstruction. The most common methods to remove twin images by experimental means, such as off-axis holography<sup>42</sup> and phase-shifting<sup>43</sup>, are not applicable in this case. Instead, we use a computational guarantee in the form of compressive sensing toward a twin-image-free reconstruction.

For the finite size in the real system, the reconstructed image point is no longer a delta function. By introducing the aperture function  $A(r) = \text{circ}(r/R)$  into the integral, the CIR of the imaging system after eliminating the twin image is calculated by setting  $\mathbf{r}_k=0$  and  $I_k=1$  in the second term of Eq. (12), which is

$$\begin{aligned} I_{\text{PSF}}(r_o) &= \iint \exp\left[\frac{i\pi}{r_1^2} (|\mathbf{r} - \mathbf{r}_o|^2 - |\mathbf{r}|^2)\right] A(|\mathbf{r}|) dS \\ &= \exp\left(\frac{i\pi}{r_1^2} r_o^2\right) \frac{R}{r_o} J_1(2\pi r_o R / r_1^2) \end{aligned} \tag{13}$$

**Reconstruction algorithm**

The captured image may be represented as the convolution of the ideal image and FZA shadow by rewriting Eq. (10) as

$$I(x, y) = O(x, y) * T(x, y) + e(x, y) \tag{14}$$

where “\*” denotes the convolution.  $O(x, y)$  is the image to be restored on the sensor plane.  $e(x, y)$  is a random term that includes photodetector noise, crosstalk, quantization noise and artifacts caused by diffraction. If we divide the cosine term of  $T(x, y)$  into  $[h(x, y) + h^*(x, y)]/2$ , where  $h(x, y) = \exp[i(\pi/r_1^2)(x^2 + y^2)]$  and  $h^*$  is the conjugate of  $h$  to express  $T(x, y)$ , Eq. (14) becomes

$$\begin{aligned} I(x, y) &= C + \frac{1}{4}[O(x, y) \times h(x, y) \\ &\quad + O(x, y) * h^*(x, y)] + e(x, y) \\ &= C + \frac{1}{4}U(x, y) + \frac{1}{4}U^*(x, y) + e(x, y) \\ &= C + \frac{1}{2}\text{Re}\{U(x, y)\} + e(x, y) \end{aligned} \tag{15}$$

where  $C = O(x, y) \times (1/2)$  is a constant. Here,  $h(x, y)$  has the same form as the Fresnel propagation kernel  $\exp[i(\pi/\lambda d)(x^2 + y^2)]$  when  $r_1^2 = \lambda d$ .  $U(x, y)$  can be regarded as the diffracted wavefront of propagating at the virtual distance of  $d$  and the virtual wavelength of  $\lambda$ .  $U^*(x, y)$  is the conjugate wave of  $U(x, y)$ . Equation (15) indicates that the measurement has the same form as the inline hologram except for the background intensity  $|U|^2$ .

Let us denote by  $N_x \times N_y = N_{xy}$  the number of im  $O(x, y)$  age samples. The constant term in Eq. (15) can be removed by filtering out the direct current component. Then, the measured image  $I \in \mathbb{R}^{N_{xy}}$  is expressed as a function related to  $O \in \mathbb{R}^{N_{xy}}$ , which is the forward transform model:

$$I = \frac{1}{2}\text{Re}\{\mathcal{F}^{-1}\mathcal{H}\mathcal{F}O\} + e \tag{16}$$

where  $\mathcal{F}$  and  $\mathcal{F}^{-1}$  are the Fourier transform operator and inverse Fourier transform operator, respectively, and  $\mathcal{H}$  indicates the operator that multiplies by the transfer function  $H = i \exp[-i\pi\lambda z(u^2 + v^2)]$ . Since  $O$  is a real function and  $H$  is a central symmetry function, Eq. (16) can be written as

$$I = \frac{1}{2}\mathcal{F}^{-1}\mathcal{H}_T\mathcal{F}O + e \tag{17}$$

where  $\mathcal{H}_T$  indicates the operator multiplied by  $H_T = \text{Re}\{H\} = \sin[\pi r_1^2(u^2 + v^2)]$ , which is the normalized Fourier transform of  $\cos[\pi(x^2 + y^2)/r_1^2]$ . Equation (17) is the representation of Eq. (14) in the frequency domain. Let  $F \in \mathbb{C}^{N_{xy} \times N_{xy}}$  be a 2D discrete Fourier transform matrix. Denote  $\Sigma$  as a diagonal matrix whose nonzero entries are the discrete value of  $H_T$ . Then, the observation  $I$  is given by  $I=KO$ , where  $K$  is

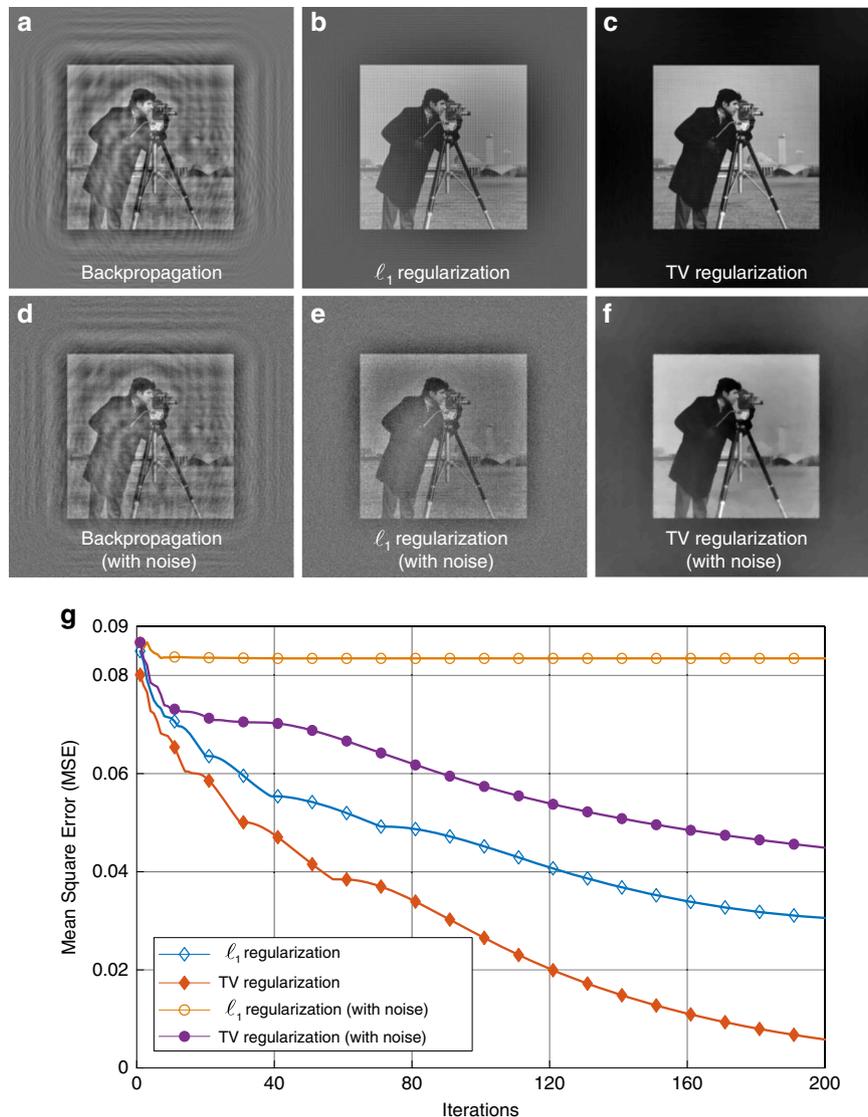
$$K = F^*\Sigma F \tag{18}$$

Solving  $O$  with a known  $I$  and forward transform  $K$  is a typical inverse problem. The solution is not unique because any value can be freely assigned to the imaginary part of  $U(x, y)$  in Eq. (15). To remedy this issue, a priori knowledge as a regularization item should be integrated into the process of image reconstruction to obtain a stable solution. For natural images, the gradient distribution tends to be zero, and object  $O$  can be regarded as sparse in the gradient domain. In contrast, the twin image generates a diffuse pattern, which is nonsparse in the gradient domain. Thus, the sparsity constraint can eliminate the twin image.

Therefore, the reconstruction can be realized by minimizing the objective function:

$$\hat{O} = \arg \min_O \frac{1}{2}\|I - KO\|_2^2 + \tau\Phi \tag{19}$$

where  $\Phi$  is the regularizer imposing the sparsity constraint and  $\|\cdot\|_2$  denotes the  $\ell_2$  norm. The regularization parameter  $\tau$  controls the relative weight of the two terms. To guarantee accurate reconstruction, the observation matrix should be subject to a restricted isometry



**Fig. 7** Image reconstruction simulation with and without noise for a grayscale image. **a** Backpropagation reconstruction. **b, c** Reconstruction using and TV regularization from a noise-free observation after 200 iterations. **d** Backpropagation reconstruction from a Gaussian-noised image. **e, f** Reconstruction using and TV regularization from a Gaussian-noised image after 200 iterations. **g** MSE vs. iteration order. See also Movie 1

property (RIP) condition according to compressive sensing theory<sup>44</sup>. Note that  $K$  is a block-circulant matrix, which is under the RIP condition with very high probability<sup>45</sup>. With a suitable regularizer, the objective function can converge rapidly and acquire a good result. Since  $K$  is related to the reconstruction distance, the method can realize numerical focusing for 3D scenes. The axial resolution is analyzed in the supplementary section (Section S2).

**Total variation denoising**

Exploiting the sparsity of the unknown image can significantly enhance reconstruction performance. The  $\ell_1$

norm and TV are widely used regularization methods. For natural scenes, TV regularization usually works better than  $\ell_1$  regularization. The TV of an image is given by the sum of magnitudes of the image gradients as

$$\|O\|_{TV} = \sum_i^{N_{xy}} \sqrt{|\Delta_i^h O|^2 + |\Delta_i^v O|^2} \tag{20}$$

where  $\Delta_i^h$  and  $\Delta_i^v$  denote the horizontal and vertical first-order local difference operations, respectively. Since the in-focus object has sharp edges while the out-of-focus twin image is diffuse, the TV of the in-focus image is much less than that of the twin image. The reconstruction

can be represented in the form of the following TV minimization problem:

$$\hat{O} = \arg \min_O \frac{1}{2} \|I - KO\|_2^2 + \tau \|O\|_{TV} \quad (21)$$

We adapt the two-step iterative shrinkage/thresholding (TwIST) algorithm<sup>46</sup> to solve Eq. (21). The TwIST algorithm is a nonlinear two-step iterative version of the iterative shrinkage/thresholding (IST) algorithm to improve the convergence rate. The TwIST algorithm provides  $\ell_1$  regularization by default. It can also be extended to TV regularization by passing a function handle. The corresponding code can be found in GitHub<sup>47</sup>. The performance of TV regularization and  $\ell_1$  regularization is shown in Fig. 7. In this case,  $z_1 = 200$  mm,  $z_2 = 3$  mm, and the size of the object is 200 mm  $\times$  200 mm. The mean square error (MSE) is introduced to evaluate reconstruction quality quantitatively. The MSE of the reconstruction error is defined as

$$\text{MSE} = \frac{1}{N_{xy}} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} [O(i, j) - \hat{O}(i, j)]^2 \quad (22)$$

Figure 7a shows the BP reconstruction calculated by Eq. (12); the original image is immersed in the noise caused by the twin image. Figure 7b, c shows the reconstructions by  $\ell_1$  and TV regularization, respectively. Both methods are able to eliminate the twin image effectively, but the TV regularization result has higher contrast and smaller MSE than the  $\ell_1$  regularization result. Moreover, TV regularization not only eliminates the twin image but also suppresses noise.

To evaluate the noise immunity capability, zero-mean Gaussian noise with 0.01 variance is added to the observed image. The BP reconstruction, the reconstruction by  $\ell_1$  and TV regularization from a noised observed image are shown in Fig. 7d–f, respectively. The  $\ell_1$  regularization result is not satisfactory due to residual noise. By contrast, TV denoising is remarkably effective in smoothing away noise while preserving edges. These results are obtained with 200 iterations. Figure 7g shows that the MSE varies with the iterations (see also Movie 1 for details). Reconstruction with TV regularization is used in our experiments.

### Fabrication

The FZA mask is fabricated on soda-lime glass substrates with a thickness of 2 mm by applying the laser direct writing technique. A Cr layer (140 nm) is deposited on the substrates by vacuum evaporation. The surface of

the Cr layer is oxidized for antireflection. A photoresist layer (1  $\mu\text{m}$ ) is then spin-coated on the Cr layer. This blank photomask is exposed under a scanning laser beam to create a latent image in the photoresist layer. After exposure, unexposed parts of the photoresist can be removed in a developer (5–7% NaOH solution). Subsequently, by immersion in an etchant ( $(\text{NH}_4)_2\text{Ce}(\text{NO}_3)_6 + \text{HClO}_4$  solution), the open area of the Cr layer is etched away to form a transparent zone, while the Cr layer protected by the photoresist is not etched to form an opaque zone. Finally, the remaining photoresist is stripped by immersion in a high concentration of developer.

### Acknowledgements

This work is supported by the National Key R&D Program of China (No. 2017YFF0106400) and the National Natural Science Foundation of China (No. 61827825).

### Author contributions

J.W., L.C., and G.B. designed and conducted the experiments and programmed the algorithm. H.Z., W.Z., and G.J. contributed to the experiments and the conducted data analyses. J.W., L.C., and G.B. wrote the manuscript.

### Conflict of interest

The authors declare that they have no conflict of interest.

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41377-020-0289-9>.

Received: 19 August 2019 Revised: 12 March 2020 Accepted: 15 March 2020

Published online: 07 April 2020

### References

- Sun, B. et al. 3D computational imaging with single-pixel detectors. *Science* **340**, 844–847 (2013).
- Xiao, X. et al. Advances in three-dimensional integral imaging: sensing, display, and applications [Invited]. *Appl. Opt.* **52**, 546–560 (2013).
- Antipa, N. et al. Single-shot diffuser-encoded light field imaging. *Proc. of 2016 IEEE International Conference on Computational Photography*. 1–11 (Evanston, IL, USA: IEEE, 2016).
- Cossairt, O. S., Miao, D. & Nayar, S. K. Gigapixel computational imaging. *Proc. 2011 IEEE International Conference on Computational Photography*. 1–8 (Pittsburgh, PA, USA: IEEE, 2011).
- Brady, D. J. et al. Multiscale gigapixel photography. *Nature* **486**, 386–389 (2012).
- Shapiro, J. H. Computational ghost imaging. *Phys. Rev. A* **78**, 061802 (2008).
- Altmann, Y. et al. Quantum-inspired computational imaging. *Science* **361**, eaat2298 (2018).
- Saunders, C., Murray-Bruce, J. & Goyal, V. K. Computational periscopy with an ordinary digital camera. *Nature* **565**, 472–475 (2019).
- Bishara, W. et al. Lensfree on-chip microscopy over a wide field-of-view using pixel super-resolution. *Opt. Express* **18**, 11181–11191 (2010).
- Greenbaum, A. et al. Imaging without lenses: achievements and remaining challenges of wide-field on-chip microscopy. *Nat. Methods* **9**, 889–895 (2012).
- Chapman, H. N. & Nugent, K. A. Coherent lensless X-ray imaging. *Nat. Photonics* **4**, 833–839 (2010).
- Witte, S. et al. Lensless diffractive imaging with ultra-broadband table-top sources: from infrared to extreme-ultraviolet wavelengths. *Light Sci. Appl.* **3**, e163 (2014).
- Sinha, A. et al. Lensless computational imaging through deep learning. *Optica* **4**, 1117–1125 (2017).
- Ren, Z. B., Xu, Z. M. & Lam, E. Y. M. End-to-end deep learning framework for digital holographic reconstruction. *Adv. Photonics* **1**, 016004 (2019).

15. Rivenson, Y., Wu, Y. C. & Ozcan, A. Deep learning in holography and coherent imaging. *Light Sci. Appl.* **8**, 85 (2019).
16. Fenimore, E. E. & Cannon, T. M. Coded aperture imaging with uniformly redundant arrays. *Appl. Opt.* **17**, 337–347 (1978).
17. Gottesman, S. R. & Fenimore, E. E. New family of binary arrays for coded aperture imaging. *Appl. Opt.* **28**, 4344–4352 (1989).
18. Stork, D. G. & Gill, P. R. Optical, mathematical, and computational foundations of lensless ultra-miniature diffractive imagers and sensors. *Int. J. Adv. Syst. Meas.* **7**, 201–208 (2014).
19. Gill, P. R. et al. Thermal escher sensors: pixel-efficient lensless imagers based on tiled optics. *Proc. Computational Optical Sensing and Imaging 2017*. (San Francisco: Optical Society of America, 2017).
20. Asif, M. S. et al. FlatCam: thin, lensless cameras using coded aperture and computation. *IEEE Trans. Comput. Imaging* **3**, 384–397 (2017).
21. DeWeert, M. J. & Farm, B. P. Lensless coded-aperture imaging with separable Doubly-Toeplitz masks. *Optical Eng.* **54**, 023102 (2015).
22. Antipa, N. et al. DiffuserCam: lensless single-exposure 3D imaging. *Optica* **5**, 1–9 (2018).
23. Gabor, D. A new microscopic principle. *Nature* **161**, 777–778 (1948).
24. Rogers, G. L. Gabor diffraction microscopy: the hologram as a generalized zone-plate. *Nature* **166**, 237 (1950).
25. Mertz, L. & Young, N. O. Fresnel transformation of images. In *Proc. ICO Conference on Optical Instruments and Techniques* (ed. K. J. Habbell) 305–310 (Chapman and Hall, London, 1962).
26. Shimano, T. et al. Lensless light-field imaging with Fresnel zone aperture: quasi-coherent coding. *Appl. Opt.* **57**, 2841–2850 (2018).
27. Nakamura, Y. et al. Lensless light-field Imaging with Fresnel Zone Aperture. In *ITE Technical Report Information Sensing Technologies (IST)*. 7–8 (The Institute of Image Information and Television Engineers, 2016).
28. Tajima, K. et al. Lensless light-field imaging with multi-phased Fresnel zone aperture. *Proc. 2017 IEEE International Conference on Computational Photography*. 1–7 (Stanford, CA, USA: IEEE, 2017).
29. Sao, M. Y. S. et al. Lensless close-up imaging with Fresnel zone aperture. *Jpn. J. Appl. Phys.* **57**, 09SB05 (2018).
30. Donoho, D. L. Compressed sensing. *IEEE Trans. Inf. Theory* **52**, 1289–1306 (2006).
31. Duarte, M. F. et al. Single-pixel imaging via compressive sampling. *IEEE Signal Process. Mag.* **25**, 83–91 (2008).
32. Huang, G. et al. Lensless imaging by compressive sensing. *Proc. 2013 IEEE International Conference on Image Processing*. 2101–2105 (Melbourne, VIC, Australia: IEEE, 2013).
33. Satat, G., Tancik, M. & Raskar, R. Lensless imaging with compressive ultrafast sensing. *IEEE Trans. Comput. Imaging* **3**, 398–407 (2017).
34. Brady, D. J. et al. Compressive holography. *Opt. Express* **17**, 13040–13049 (2009).
35. Rivenson, Y., Stern, A. & Javidi, B. Overview of compressive sensing techniques applied in holography [Invited]. *Appl. Opt.* **52**, A423–A432 (2013).
36. Rivenson, Y., Stern, A. & Javidi, B. Improved depth resolution by single-exposure in-line compressive holography. *Appl. Opt.* **52**, A223–A231 (2013).
37. Stern, A. *Optical Compressive Imaging*, Ch. 8 (CRC Press, Boca Raton, 2016).
38. Zhang, W. H. et al. Twin-image-free holography: a compressive sensing approach. *Phys. Rev. Lett.* **121**, 093902 (2018).
39. Beynon, T. D., Kirk, I. & Mathews, T. R. Gabor zone plate with binary transmittance values. *Opt. Lett.* **17**, 544–546 (1992).
40. Beynon, T. D. & Strange, R. M. R. Computational study of diffraction patterns for near-field Fresnel and Gabor zone plates. *J. Opt. Soc. Am. A* **17**, 101–106 (2000).
41. Salgado-Remacha, F. J., Sanchez-Brea, L. M. & Bernabeu, E. Effect of fill-factor on the Talbot effect of diffraction gratings. *J. Eur. Optical Soc. Rapid Publ.* **6**, 11055 (2011).
42. Leith, E. N. & Upatnieks, J. Reconstructed wavefronts and communication theory. *J. Opt. Soc. Am.* **52**, 1123–1130 (1962).
43. Yamaguchi, I. & Zhang, T. Phase-shifting digital holography. *Opt. Lett.* **22**, 1268–1270 (1997).
44. Candes, E. J. & Tao, T. Decoding by linear programming. *IEEE Trans. Inf. Theory* **51**, 4203–4215 (2005).
45. Marcia, R. F. & Willett, R. M. Compressive coded aperture superresolution image reconstruction. *Proc. 2008 IEEE International Conference on Acoustics, Speech and Signal Processing*. 833–836 (Las Vegas, NV, USA: IEEE, 2008).
46. Bioucas-Dias, J. M. & Figueiredo, M. A. A new TwIST: two-step iterative shrinkage/thresholding algorithms for image restoration. *IEEE Trans. Image Process.* **16**, 2992–3004 (2007).
47. Wu, J. & Cao L. Sourcecode website: [https://github.com/THUHoloLab/FZA\\_lensless\\_imaging](https://github.com/THUHoloLab/FZA_lensless_imaging) (2019).