PHOTONICS Research

Fast X-ray ghost imaging by scanning aggregate patterns

Jin-Tao Xie,[1,2] **Hui Zeng,**[3,4] **Ming-Fei Li,**[1,2,5] **Yi-Fei Li,**[1] **Jin-Guang Wang,**[1] **Bao-Gang Quan,**[1,2] **Li-Ming Chen,**[3,4,6] **and Ling-An Wu**[1,2,*]

[1]*Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China*
[2]*School of Physical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China*
[3]*School of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai 200240, China*
[4]*IFSA Collaborative Innovation Center, Shanghai Jiao Tong University, Shanghai 200240, China*
[5]*e-mail: mf_li@iphy.ac.cn*
[6]*e-mail: lmchen@sjtu.edu.cn*
**Corresponding author: wula@iphy.ac.cn*

Received 25 February 2025; revised 4 July 2025; accepted 6 July 2025; posted 7 July 2025 (Doc. ID 560065); published 5 September 2025

For practical applications of X-ray ghost imaging (XGI), the imaging time is a major challenge. In this paper, we propose a fast XGI scheme based on a continuous translation mask with etched aggregate patterns. High contrast-to-noise ratio images are obtained with an exposure time of only 4 s and 24 s for a $3.4\ \mathrm{mm} \times 3.8\ \mathrm{mm}$ and $5.9\ \mathrm{mm} \times 6.1\ \mathrm{mm}$ field-of-view, respectively. The spatial resolution can reach $\sim 150\ \mu\mathrm{m}$. The influences of the sampling frequency, the mask scanning speed, and the detector integration time on image quality are examined, from which we demonstrate that the imaging time can be further reduced by increasing the mask translation speed. By applying a compressed sensing reconstruction algorithm, the exposure time is greatly reduced while maintaining image quality. Our method indicates a path for the development of future XGI applications.

https://doi.org/10.1364/PRJ.560065

1. INTRODUCTION

Ghost imaging (GI) is an indirect imaging technique that reconstructs the image of a sample/target via the second-order correlation [1] between the spatially modulated patterns illuminating the object and the total transmitted or reflected intensity, which is recorded by only a single-pixel "bucket" detector. Its first demonstration can be traced back to 1995, when a two-photon quantum source was used to produce the correlation [2]. One of the photons interacted with an object and was collected by a photon counting "bucket" detector, while the other without interaction was recorded by an x–y scanning fiber single-photon detector. After many exposures, a magnified image of the object was observed in the coincidence counting rate. This kind of nonlocal imaging, a seemingly "spooky" action from quantum entangled photons, gave rise to the name of GI [2]. However, since then, most experiments have been performed with classical light sources, which are more suitable for practical applications, and have already covered the spectrum from microwave [3] to terahertz [4–9] to hard X-rays [10–13], and have even exploited atoms [14], electrons [15], and neutrons [16–18]. The great appeal of GI is its ability to reconstruct images under very low illumination and even through turbulent media or at single-photon levels [12,19,20], while spatial resolution is not limited by Rayleigh diffraction [21]. In most visible light setups, the illumination patterns were created by a computer-controlled digital micromirror device (DMD) or a spatial light modulator (SLM), and this single beam path scheme was called computational GI [22], or single-pixel imaging (SPI) [23].

X-ray analysis is well known in clinical medicine and nondestructive inspection because of its high penetration. For X-ray GI (XGI), the main challenge is how to create a series of intensity modulated two-dimensional patterns to illuminate the object and then record the total transmitted intensity by a bucket detector. The spatial modulators used in visible and infrared optics are generally not suitable for the direct modulation of X-rays, and we have to employ some high-absorption or high-refraction materials to create intensity speckles. Modulators that have been used include silicon carbide sandpaper [12], copper [13], gold [13], copy paper [24], and glass or metal powder [25,26]. However, these schemes required prerecording of the patterns and such large amounts of data had to be stored and processed, all taking much more time. With regard to the original GI schemes based on a beamsplitter, for X-rays a crystal has been used, but this requires coherent radiation as from a synchrotron [25,26] or the addition of a monochromator [24], and diffraction inevitably involves losses.

Fast indirect modulation has been demonstrated by the spatial modulation of a Ti:sapphire femtosecond laser beam used to pattern the reflection of an X-ray beam from a crystal [27], but again, this required both beams to be coherent.

Various extensions of XGI, such as computer tomography [26], phase-contrast-based imaging [28,29], and element differentiation based on fluorescence and absorption spectra [30,31] have also been reported. However, the total time required to expose all the measurement patterns one by one is the greatest problem. The low modulation frequency results in long imaging time in existing XGI setups, which can be up to several hours [10–13,24–26,28–31] and is therefore impractical for real applications compared to conventional X-ray imaging methods. Our motivation is to reduce the imaging time of XGI to meet the fast-imaging requirements in real world scenarios. For this, the sampling patterns of previous XGI setups can be optimized. The optimal patterns should have the following three characteristics. (1) They can be contracted to an aggregate pattern of the smallest possible size. (2) They can be scanned as fast as possible. (3) The computing time for image reconstruction should be as short as possible.

Inspired by compressed ultrahigh-speed single-pixel imaging based on swept aggregate patterns [32], which combines a DMD with laser scanning hardware to achieve pattern projection rates of up to 14 MHz and tunable frame sizes of up to 101 × 103 pixels, we propose a method that can greatly shorten the imaging time for XGI. It should be noted that the DMD and the rotating polygonal mirror used for ultrahigh-speed modulation in Ref. [32] are not suitable for X-rays that would just pass through. By reshaping and continuously scanning a cyclic S-matrix modified into an aggregate pattern, it is possible to collect all the bucket signals very rapidly from a continuous voltage-time signal recorded by an inexpensive single-pixel scintillator X-ray detector. To demonstrate the feasibility of our scheme, we performed a proof-of-principle experiment employing a mask with the designed aggregate patterns fabricated on an inexpensive printed circuit board (PCB), and we reconstructed high contrast-to-noise ratio images of an object using two mask sizes with imaging time as short as 4 s. The spatial resolution can reach ∼150 μm, which is comparable to the pixel size of the flat panel X-ray detectors on the market and sufficient for chest radiographic imaging [33]. The total cost of our modulation mask and X-ray single-pixel detector is about an order of magnitude lower than that of an X-ray flat panel detector. Compared to the method in Ref. [27], our setup with only a fabricated mask and two translation stages is much simpler and lower in cost. We also point out the capability of our setup for further reducing the imaging time, and we believe that the modulation frequency will not be limited to 10 kHz as technology improves. Although the frame size of our current mask is still rather small, it should not be difficult to find some way to increase it in the near future, such as by multiple translation scanning and using a few-pixel detector [16,34].

In Section 2 we will introduce our proposed scheme, including the setup, the aggregate pattern design, and the reconstruction algorithms. The reason for choosing the cyclic S-matrix and the specific experimental procedure will be shown in detail. The simulation and experimental results are analyzed in Section 3. We show that there is a trade-off between the degree of motion blurring and signal intensity. The quality of the reconstructed image is influenced by the integration time of a single measurement, i.e., the choice of the bucket intensity for each modulation pattern. Finally, the paper is concluded in Section 4.

2. METHOD

A. System Setup

A schematic of our system is shown in Fig. 1. An X-ray tube (Incoatec Source IusCu) operating at 45 kV with an adjustable current is used as the light source. An aperture is placed after it to ensure that only one matrix pattern is illuminated in each

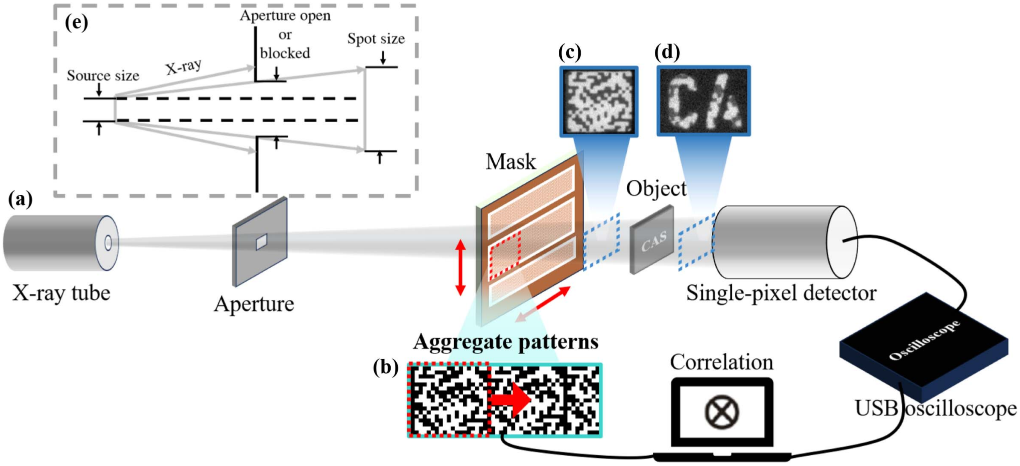

Fig. 1. Experimental setup. (a) A polychromatic X-ray beam passes through a rectangular aperture and the aggregate patterned mask to illuminate the object. The transmitted intensity is recorded by a single-pixel detector, and its output voltage is conveyed to an oscilloscope as a function of time and then displayed on a computer. (b) Close-up of a typical aggregate pattern sequence of frame size 17 × 19 pixels. The red dotted box is the section illuminated by the X-ray beam. The red arrow denotes the scanning direction. (c) and (d) Actual intensity distributions of the beam just behind the mask and behind the object (in the locations indicated by the two blue boxes) recorded by an X-ray sCMOS camera, respectively. (e) Diagram of the geometry of the X-ray beam propagation.

measurement. Two tungsten apertures, 2 mm thick and 2.89 mm × 3.23 mm and 4.72 mm × 4.88 mm in size, are used to collimate the beam onto frames of 17 × 19 and 59 × 61 pixels on the mask plate, the width of the pixels being 200 μm and 100 μm, respectively. For these two frame sets, the corresponding best spatial resolutions of the reconstructed images are therefore, theoretically, also 200 μm and 100 μm [22,35], while the fields-of-view (FOVs) are 3.4 mm × 3.8 mm and 5.9 mm × 6.1 mm, respectively. The designed aggregate patterns are etched on a mask plate composed of a 3-oz-(≈105 μm) thick copper foil on a 500-μm-thick laminate substrate. The fabrication was outsourced to a PCB manufacturing company. The length-to-width ratio of the apertures is chosen to equal that of the modulation pattern to ensure that only one pattern is projected onto the object. The intensity distributions of the beam immediately behind the mask and behind the object, as recorded by an X-ray sCMOS camera, can be seen in Figs. 1(c) and 1(d), respectively. Any pattern blur caused by diffraction can be ignored because the wavelength of the X-ray (0.01–10 nm) is much smaller than the pixel width of the mask (100 μm). The magnification of the spatial field distribution with the increase of distance is basically just a linear geometric relationship, as shown in Fig. 1(e). The object is set behind the mask as close as possible (1.5 cm) to minimize any blurring of the modulation patterns caused by the uncollimated beam. The mask plate is mounted on an x–y translation motorized stage (Sigmakoki OSMS26-100) with a 5-phase-stepping motor controlled by a two-axis controller (Sigmakoki SHOT-302GS) with 250 micro-steps so that the aggregate patterns were scanned with high precision with no noticeable mechanical vibration. The settling time of the stepping motor is not specifically given in the product specifications, but we set the shortest scanning time to 150 frames/s, i.e., each frame is exposed for about 7 ms, while the motor takes ∼1 μm to complete a pulse, so the settling time can be neglected relative to the exposure time.

The single-pixel detector (NNVT M3021-01) is a common X-ray integration detector composed of a sodium iodide scintillator, a photomultiplier tube (PMT), a high voltage module, and a preamplifier. The effective detection diameter is 25 mm, and the voltage is set to 720 V. The output from the detector enters the USB oscilloscope (TiePie Handyscope HS6), which transmits the voltage signals as a function of time to be displayed on the computer. After correlating the aggregate patterns and corresponding voltage signals, the object image can be reconstructed. If the sampling frequency of the oscilloscope is sufficiently high and the reconstruction algorithm is unchanged, the imaging speed only depends on the mechanical translation speed of the stage and the length of the aggregate patterns (which depends on the required FOV and image resolution).

B. Aggregate Pattern Designs

The aggregate patterns are created by reshaping cyclic S-matrices, which have been demonstrated to be advantageous for high-speed single-pixel imaging [32]. The rows of an S-matrix are pseudo-random sequences, or codewords in a simplex code (hence the name) [36]. To the best of our knowledge, there are three methods that are used to construct cyclic S-matrices [36]. To realize a rectangular FOV, the twin prime construction method is the most suitable. This construction produces cyclic S-matrix $\mathbf{S}$ of order n whenever $n = p(p+2)$, and both p and $p+2$ are prime numbers. Let $S_{i,j}$ denote the element of $\mathbf{S}$ with row index i and column index j. The initial row $S_{0,j}(j = 0, \cdots, n-1)$ can determine all elements of $\mathbf{S}$ according to

$$S_{i,j} = \begin{cases} S_{0,i+j}, & \text{if } i+j < n \\ S_{0,i+j-n}, & \text{if } i+j > n \end{cases}, \tag{1}$$

where $i = 0, \ldots, n-1$. An integer x is regarded as a quadratic residue (mod y) if $x \neq 0$ (mod y), and there exists another integer z such that $x \equiv z^2$ (mod y). In this case, suppose p and $q = p + 2$ are both prime numbers, such as $p = 3$ and $q = 5$, so we can define

$$f(j) = \begin{cases} +1, & \text{if } j \text{ is a quadratic residue } (\text{mod}\, p) \\ 0, & \text{if } j \equiv 0 \ (\text{mod}\, p) \\ -1, & \text{otherwise} \end{cases}, \tag{2}$$

$$g(j) = \begin{cases} +1, & \text{if } j \text{ is a quadratic residue } (\text{mod}\, q) \\ 0, & \text{if } j \equiv 0 \ (\text{mod}\, q) \\ -1, & \text{otherwise} \end{cases}, \tag{3}$$

$$S_{0,j} = \begin{cases} 0, & \text{if } [f(j) - g(j)]g(j) = 0 \\ +1, & \text{otherwise} \end{cases}. \tag{4}$$

A typical process for creating an aggregate pattern is shown in Fig. 2. A cyclic S-matrix created by the twin prime construction for $p = 3$ and $q = 5$ is depicted in Fig. 2(a), and its reshaped rows are shown in Fig. 2(b). The first three rows and patterns are marked by red, green, and blue boxes, respectively. We can see that adjacent patterns are very similar, and only the last column of the latter pattern differs from the original pattern. Thus, we can contract these patterns to the smallest size and obtain the aggregate pattern of Fig. 2(c). Every time the scanning area moves a pixel to the right, a new pattern is illuminated. For a mechanical translation speed of v and pixel width of l, we can write the total exposure time t as

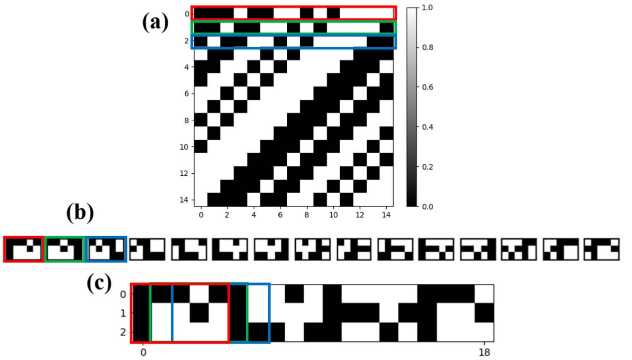

Fig. 2. (a) Cyclic S-matrix of size 15 × 15 pixels. (b) New patterns of size 3 × 5 pixels reshaped from each row. (c) Aggregate pattern of size 3 × 19 pixels.

$$t = \frac{n \times l}{v}. \tag{5}$$

Note that the required length of the aggregate pattern L is equal to $l(q + n - 1)$. For a 17 × 19 pixels frame and 200 μm pixel width, we would require a 6.82 cm long matrix, which can be directly etched on a mask of size 12 cm × 12 cm. For a 59 × 61 pixels frame and 100 μm pixel width, the length would be 36.59 cm, so we had to split this into five sections.

It should be noted that the patterns of the cyclic S-matrix are orthogonal under translation, so we can obtain full information of the object after scanning the aggregate patterns [32,37]. There are also other matrices, such as uniformly redundant arrays (URAs) [38,39] and random matrices [40,41], that can be contracted to an aggregate pattern of the smallest size. The patterns of the former are also orthogonal under translation, while those of the latter are not. Nonetheless, the latter has the advantage of decimation when applying compressed sensing (CS) [42,43]. To explain why we chose the cyclic S-matrix to fabricate the mask, we present the three aggregate patterns, their autocorrelation plots, and the images reconstructed by traditional GI and CS in Fig. 3. The URA patterns can be contracted to an aggregate pattern with a square shape, and three of them are outlined in the red, green, and blue dotted boxes shown in Fig. 3(a). To scan all the patterns with minimal translation, the zig-zag order should be applied, which means that the translation direction has to be changed accordingly. Thus, the scanning speed of the URA cannot be faster than that of the random matrix and cyclic S-matrix whose translation direction remains unchanged, as shown in Figs. 3(b) and 3(c). The autocorrelation plots of the three aggregate patterns in Figs. 3(a1), 3(b1), and 3(c1) verify that the URA and cyclic S-matrices are orthogonal, while the random matrix is not. Due to the orthogonality of the former two matrices, the image of the simulated letter "C" can be perfectly reconstructed within 1 ms at full sampling with the traditional GI algorithm, as shown in Figs. 3(a2) and 3(c2), while it cannot be reconstructed by the random matrix in Fig. 3(b2). The random matrix can reconstruct a high-quality image using a CS algorithm (here, we used the augmented Lagrangian and alternating direction based on the total variation algorithm (TVAL3) [44]), but the computation takes 340.4 ms because it needs many iterations, as depicted in Fig. 3(b3). The CS algorithm can also be applied to the other two matrices in Figs. 3(a3) and 3(c3), but the computing times are still long. In view of the demands of future real-time XGI setups, we chose the cyclic S-matrix for the design of our mask.

C. Reconstruction Algorithms

1. Single-Pixel Imaging

In the SPI framework, the imaging process can be expressed as

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \tag{6}$$

where $\mathbf{A} = [\mathbf{a}_1, \cdots, \mathbf{a}_n]^T$ is the modulation matrix, where $\mathbf{a}_i$ $(i = 1, \ldots, n)$ is an n-element row vector, $\mathbf{x} = [x_1, \cdots, x_n]^T$ denotes the object, and $\mathbf{b} = [b_1, \cdots, b_n]^T$ represents the single-pixel detector signal. The general reconstruction expression to obtain the object information is

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}, \tag{7}$$

where $\mathbf{A}^{-1}$ denotes the inverse of $\mathbf{A}$.

2. Ghost Imaging

In the traditional GI framework, the ghost image $G(x, y)$ can be retrieved after n measurements by calculating the correlation using [12]

$$\begin{aligned} G(x,y) &= \langle BI(x,y)\rangle - \langle B\rangle\langle I(x,y)\rangle \\ &= \frac{1}{n}\sum_{i=1}^{N} B_i I_i(x,y) - \frac{1}{n^2}\sum_{i=1}^{N} B_i \sum_{i=1}^{N} I_i(x,y), \end{aligned} \tag{8}$$

where B is the integrated bucket detector intensity, $I(x, y)$ is the pre-designed intensity modulation mask pattern, and $\langle\cdot\rangle$ denotes an ensemble average. Here, we assume a binary

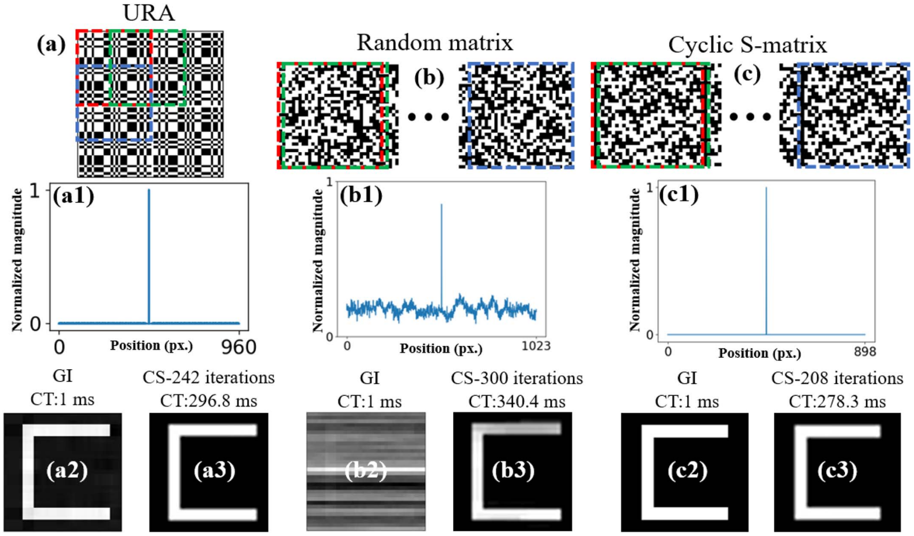

Fig. 3. (a)–(c) Aggregate patterns of the URA of size 31 × 31 pixels, random matrix of size 32 × 32, and cyclic S-matrix of size 29 × 31, respectively. (a1), (b1), (c1) Autocorrelation plots of the corresponding three matrices. (a2) and (a3), (b2) and (b3), (c2) and (c3) Images reconstructed by the above three matrices with the GI and CS algorithms, respectively. CT, computation time.

modulation with elements of value 1 or 0; then if $B_i = b_i$, $I_i(x, y)$ is the two-dimensional reshaped form of $\mathbf{a}_i$ and $\mathbf{A} = \mathbf{S}$. With the equation $\mathbf{JS} = \frac{n+1}{2}\mathbf{J}$ [36], the above expression can be described in matrix form as

$$\begin{aligned}\mathbf{G} &= \frac{1}{n}\mathbf{S}^T\mathbf{S}\mathbf{x} - \frac{1}{n^2}\|\mathbf{b}\|_1 \cdot \frac{n+1}{2}\mathbf{J} \\ &= \frac{1}{n}\mathbf{S}^T\mathbf{S}\mathbf{x} - \frac{n+1}{2n^2}\mathbf{J}\mathbf{S}\mathbf{x},\end{aligned} \tag{9}$$

where $\|\mathbf{b}\|_1$ represents the sum of all elements in vector $\mathbf{b}$, and $\mathbf{J}$ is the all-one matrix. Considering the equations $\mathbf{S}^T\mathbf{S} = \frac{n+1}{4}(\mathbf{J} + \mathbf{I})$ and $\mathbf{JS} = \frac{n+1}{2}\mathbf{J}$ [36], we can simplify Eq. (9) as

$$\mathbf{G} = \frac{n+1}{4n}\mathbf{x} - \frac{n+1}{4n^2}\mathbf{x}_{\text{sum}} = \frac{n+1}{4n}(\mathbf{x} - \mathbf{x}_{\text{mean}}), \tag{10}$$

where $\mathbf{x}_{\text{sum}} = [\|\mathbf{x}\|_1, \cdots, \|\mathbf{x}\|_1]^T$ of size $n \times 1$, and $\mathbf{x}_{\text{mean}} = \frac{\mathbf{x}_{\text{sum}}}{n}$, whose element is the mean value of the object vector. Although the ghost image vector $\mathbf{G}$ is not absolutely equal to the object vector $\mathbf{x}$ in the equation, we still can obtain full information from the object due to the self-adaptive color bar in the computer program. In most GI cases, the absolute gray value of the reconstructed image is meaningless, while the relative gray value can provide full information about the object.

Note that the value of the pre-designed mask pattern cannot be 0 in reality as it depends on the absorption of the material in the X-ray region. Thus, it is important to measure the modulation depth of our mask. The parameters of the X-ray tube were set at 45 kV and 10 μA, and the number of photons over an FOV of 5.9 mm × 6.1 mm was recorded for 20 ms by the single-pixel detector. Figure 4 shows the photon counts with the beam completely blocked (bottom red dots) and unblocked (top green crosses) by the copper mask. The background count is only influenced by occasional cosmic rays in our setup, which is negligible, so there is no background to be subtracted. From this, we can calculate the modulation depth to be $a = 91.8\%$ and the photon emission from the X-ray tube to be $P = 1.25 \times 10^5$ photons/s. It should be noted that the number of the black and white pixels of each pattern is $\frac{n-1}{2}$ and $\frac{n+1}{2}$ [36], respectively. Thus, the numbers of 0 and 1 are as nearly equal as possible with the increase of n. Because the radiation dose is very important in X-ray applications, we add the following equation to calculate this parameter in our experiment:

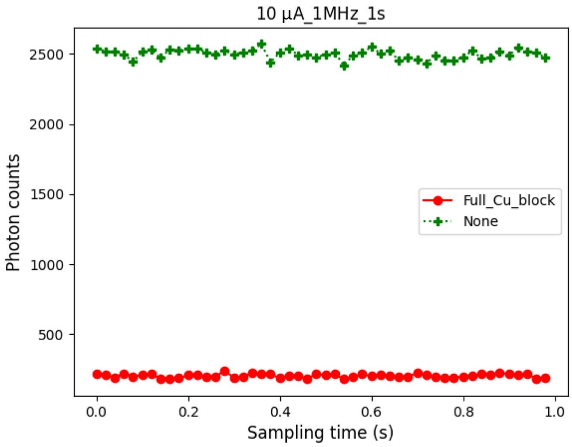

Fig. 4. Number of photons emitted from the X-ray tube blocked (bottom red dots) and unblocked (top green crosses) by the Cu mask. The detector's sampling time and frequency are 1 s and 1 MHz, respectively. Each integration time is 20 ms.

$$D = \left[\frac{P(1-a)}{2} + \frac{P}{2}\right]t, \tag{11}$$

where D denotes the radiation dose with the unit of photons, and t is the imaging time. The first part $\frac{P(1-a)}{2}$ is the photon count rate passing through the modulation area, which consists of approximately 50/50 0 and 1 bits.

3. Compressed Sensing

To improve the image quality and reduce the sampling rate, we also used CS, which excels at recovering sparse signals from subsampled data and has been validated in numerous publications [45]. Here, we adopted TVAL3 to minimize the cost function,

$$\min_{\mathbf{x}} \sum_i \|\mathbf{D}_i\mathbf{x}\|_2 + \frac{\mu}{2}\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2, \quad \text{s.t. } \mathbf{x} > 0, \tag{12}$$

where $\|\mathbf{x}\|_2$ is the l_2 norm of vector $\mathbf{x}$, $\mathbf{x} = [x_1, \cdots, x_m]^T$, $m \le n$ is the estimated object, $\mathbf{b} = [b_1, \cdots, b_m]^T$ is the bucket intensity integrated for a given time, $\mathbf{A}$ is the cyclic S-matrix of size $m \times n$, $\mu > 0$ is the penalty parameter of the quadratic fidelity term, $\mathbf{D}_i\mathbf{x}$ $(i = 1, \cdots, m)$ is the discrete gradient vector of $\mathbf{x}$ at position i, and $\sum \|\mathbf{D}_i\mathbf{x}\|_2$ is the total variation. The fidelity term $\|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2^2$ minimizes the difference between the estimation and measurements, while the total variation removes the image noise and preserves edge information at the same time [44].

4. Image Quality Assessment

To quantify the reconstruction quality, we adopt the contrast-to-noise ratio (CNR) [46] and structural similarity index (SSIM). The peak signal-to-noise ratio is not considered here because it is almost not affected by motion blur in our method. The CNR is defined as

$$\text{CNR} = \frac{\langle G_1 \rangle - \langle G_0 \rangle}{\sqrt{\sigma_1^2 + \sigma_0^2}}, \tag{13}$$

where G_1, G_0 are the reconstructed pixel values when the ground-truth object transmissions are 1 and 0 (thus limited to a binary object), respectively, and σ_1, σ_0 are their respective standard deviations.

3. RESULTS AND DISCUSSION

A. Simulation Results

Most existing XGI devices require the modulation pattern to be stationary relative to the object during exposure to obtain a bucket detector intensity signal. Thus, the total imaging time also depends on the sequential mechanical movement of the translation stages. This cannot be very short because there are many stage movements involved. Here, continuous translational scanning is employed, which greatly speeds up the exposure rate while ensuring sufficiently long enough acquisition time for all the patterns. However, as the mask moves, it is inevitable that the bucket detector intensity will contain some photons passing through the next pattern. This process can be explained in Figs. 5(a) and 5(b). Assume that the X-ray beam illuminates the ith and $(i+1)$th patterns at times t_i and t_{i+1},

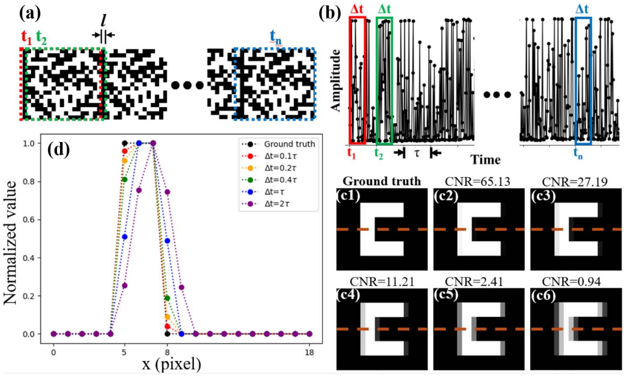

Fig. 5. Simulation results with motion blur for our setup. (a) Part of the aggregate pattern sequence for a frame size of 17 × 19 pixels. The red, green, and blue dotted boxes mark the moments t_1, t_2, and t_n at which the X-ray beam illuminates the 1st, 2nd, and nth patterns, respectively. (b) Typical bucket detector signal recorded when the mask moves from the first to the last pattern. (c1)–(c6) Ground truth of virtual object letter "C" and its reconstructed images after integration times of 0.1τ, 0.2τ, 0.4τ, τ, and 2τ. (d) Normalized intensities of the cross sections indicated by the dashed lines in (c1)–(c6).

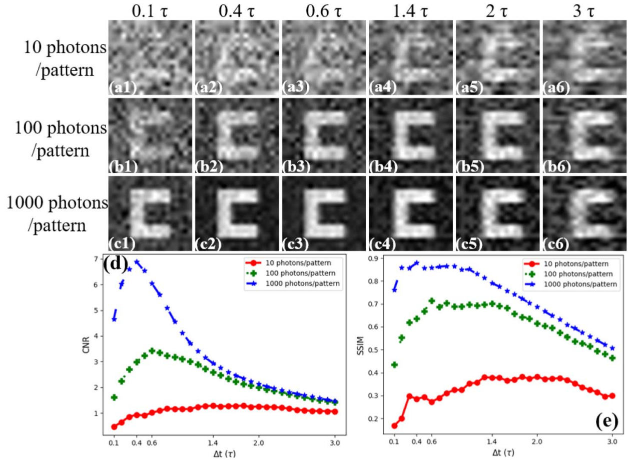

Fig. 6. Simulation results incorporating motion blur and shot noise into our setup. (a1)–(c6) Images reconstructed by a GI algorithm, (d) and (e) their CNR and SSIM curves for different integration times (0.1τ to 3τ) and photon fluxes (10–1000 photons/pattern).

respectively, and the interval between them is $t_{i+1} - t_i = \tau$. The bucket detector intensity for the ith pattern can be expressed as

$$B_i = \int_{t_i}^{t_i+\Delta t} F(t)\mathrm{d}t, \tag{14}$$

where $\Delta t = \alpha\tau$ is the integration time and α a coefficient, $F(t)$ is a form of signal function shown in Fig. 5(b), $\tau = t_{i+1} - t_i = \frac{l}{v}$ and $\frac{1}{\mathrm{d}t}$ can be regarded as the detector sampling frequencies. The image is reconstructed from all the bucket values and their corresponding patterns. To examine the effect of the motion of the mask on the reconstructed images, we first perform a simulation in which the virtual object letter "C" is reconstructed with integration times of 0.1τ, 0.2τ, 0.4τ, τ, and 2τ. The results are shown in Figs. 5(c1)–5(c6), respectively, and the image cross sections along the red dashed line are plotted in Fig. 5(d). We can see that the CNR of the reconstructed image decreases and the edges become more horizontally blurred in the pattern translation direction with the increase of the integration time. Therefore, we can conclude that for a high-resolution image it would be better to choose as short an integration time as possible to minimize the effect of motion blurring. However, in real experiments, a shorter integration time means fewer signal photons, which would result in low signal-to-noise ratio images. Thus, the optimum integration time must be chosen according to the experimental conditions. This can be done by reconstructing images with varying integration time and choosing the one giving the best quality.

To examine the influence of shot noise in addition to motion blur on the reconstruction, we perform a further simulation as follows. First, the size of the aggregate pattern 17 × 341 is increased 50 times to 850 × 17,050, while the 17 × 19 size of each pattern is increased to 850 × 950. During scanning, each time the enlarged aggregate pattern moves a pixel, a new enlarged pattern is projected onto the virtual object letter "C," and the sum of the gray scale of all the pixels is taken to be the signal value at that moment. If we take this value at the given moment as the bucket intensity, the integration time Δt is equal to $\tau/50$. This takes into consideration the effect of motion blur, and the corresponding results are shown in Fig. 5(d). To simulate the shot noise, we project a uniformly distributed random sequence ranging in value from 0 to 1 divided into 10, 100, or 1000 elements onto an all-zero matrix of size 17 × 19 pixels and then enlarge it to 850 × 950. The matrix with photon fluctuations can be regarded as the intensity distribution behind the aperture and is projected onto the enlarged pattern and the object "C." The simulation results of the shot noise and motion blur are depicted in Fig. 6. The images in the first, second, and third rows are reconstructed under different photon flux conditions with 10, 100, and 1000 photons/pattern, respectively. Those in the six columns are reconstructed with different integration times of $0.1\tau, 0.4\tau$, $0.6\tau, 1.4\tau, 2\tau$, and 3τ from left to right. Figures 6(d) and 6(e) are the CNR and SSIM values versus the integration time plotted for different photon flux conditions, respectively. By comparing the images in each column, we can see that as the photon flux increases, the image becomes clearer, and the corresponding CNR and SSIM also increase. However, by observing each row of the images, we find that although increasing the integration time introduces more photons, motion blur becomes more serious. Thus, the CNR and SSIM values will first increase and then decrease as the integration time increases. The more abundant the number of photons is, the smaller the optimal integration time will become. This means that through a suitable combination of a high intensity X-ray source, a high sensitivity detector, and high sampling rate with a fast mechanical translation motor, real-time X-ray ghost imaging will become possible.

B. Experimental Results

The results of our first experiment are shown in Figs. 7 and 8. The X-ray tube was operated at 45 kV and 300 μA, and an aperture of size 2.89 mm × 3.23 mm was used for the frame

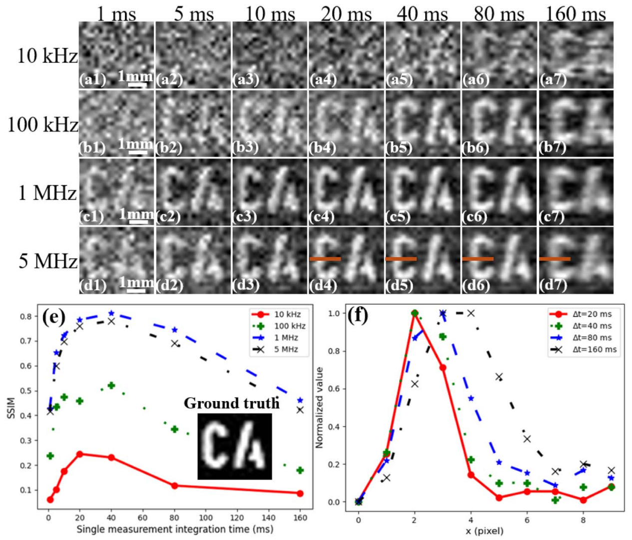

Fig. 7. Images reconstructed by GI (a1)–(d7) and their SSIM curves (e) for different detector integration times (1–160 ms) and sampling frequencies (10 kHz to 5 MHz). X-ray tube current, 300 μA; mask translation speed, 4 mm/s. (f) Normalized intensities of the cross sections indicated by the solid lines in (d4)–(d7) for different values of the integration time Δt.

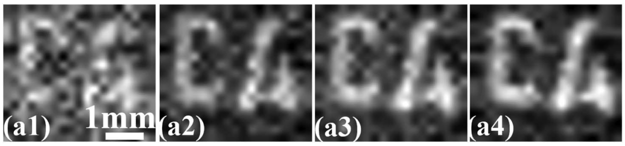

Fig. 8. Images reconstructed by GI. Tube current, 300 μA; sampling frequency, 5 MHz. Translation speed, 15 mm/s with single measurement integration times of (a1) 1 ms, (a2) 5 ms, (a3) 10 ms, and (a4) 20 ms.

size of 17 × 19 pixels with pixel width of 200 μm, so the FOV was 3.4 mm × 3.8 mm. The object was a 2-mm-thick tungsten plate with the letters "CAS" cut out. The Cu mask etched with 323 patterns was translated horizontally at a speed of $v = 4$ mm/s, and the detector sampling frequency $\frac{1}{\mathrm{d}t}$ was varied from 10 kHz to 5 MHz. The detector integration times Δt for each measurement were also varied from 1 to 160 ms. Based on the results in Fig. 6, the more photons recorded in a given short period of time, the higher the quality of the reconstructed images will be. Thus, our PMT was operated in the linear direct current mode instead of the photon counting mode. The bucket intensity for each measurement is the voltage value obtained by Eq. (14), not the number of photons. The reconstructed GI images are shown in Figs. 7(a)–7(d) and their corresponding SSIM curves in Fig. 7(e). The image inset in Fig. 7(e) is the ground truth, obtained by resizing the image of the object captured by a high resolution X-ray sCMOS camera. Because the downsampling process cannot produce a perfect binary image with only a few pixels, the post-downsampling gray image is more suitable for calculating the SSIM. The normalized intensities of the cross sections marked by the solid lines in Figs. 7(d4)–7(d7) are plotted in Fig. 7(f). We can see (from the images down along a column) that as the sampling frequency is increased from 10 kHz to 1 MHz, the SSIM of the images with the same integration time is improved, while for 1 MHz and 5 MHz there is little difference. This is because when the sampling frequency is sufficiently fast the detector can collect most of the photons, while at a lower frequency many photons will be lost because the photons in the intervals between the sampling times cannot be detected. For short integration times, the noise caused by random cosmic rays, random background photons, and the readout circuit will lead to a low SSIM image, as shown in the 1 ms column of Fig. 7. With the increase of the integration time from 1 ms to 40 ms at the frequency of 1 MHz and 5 MHz, the SSIM also increases as the signal photons become more plentiful. From Fig. 7(e), we can also see that as the sampling frequency is increased from 10 kHz to 1 MHz the SSIM increases markedly. However, at 80 ms and 160 ms integration times with sampling frequencies of 1 MHz and 5 MHz, the SSIM starts to decrease, and a distinct motion blurring effect is observed in the corresponding two columns, even though there are plenty of signal photons. Figure 7(f) shows that as the integration time Δt increases from 20 to 160 ms, the cross section of the vertical line in the letter "C" becomes broader due to blurring of the image. This agrees well with our simulation results.

Next, the mask translation speed was increased from 4 mm/s to 15 mm/s, in effect reducing the total imaging time from 16.15 s to 4.3 s, and images were obtained at a sampling frequency of 5 MHz with different integration times. Figure 8 shows that we still can clearly distinguish the object information of the images when the sampling frequency, the tube current, and the integration time are high enough, which means we could continue to reduce the imaging time by increasing the translation speed.

To expand the FOV and improve the spatial resolution, we employed the second set of modulation frames where the number of patterns was 59 × 61 and the pixel width was 100 μm. Thus, the FOV was enlarged to 5.9 mm × 6.1 mm, and the theoretical best spatial resolution improved to 100 μm [22,35]. The size of the aperture was adjusted to 4.72 mm × 4.88 mm and the experiment run at two translation speeds of 4 mm/s and 15 mm/s, with sampling frequencies of 5 MHz and 20 MHz, respectively. Images were reconstructed by both GI and TVAL3 algorithms, and the results are given in Fig. 9. In Fig. 9(e) the corresponding SSIM values are plotted as a function of the integration time, and the inset image is the ground truth used to calculate the SSIM. Figures 9(a) and 9(b) show the images retrieved for the low 4 mm/s speed (top two rows) and Figs. 9(c) and 9(d) show those retrieved for the high 15 mm/s speed (lower two rows). The corresponding imaging times are calculated to be 90 s and 24 s, respectively. We can see that the former images are of better quality than the latter. The SSIM values of Figs. 9(a1)–9(a4) and Figs. 9(c1)–9(c4) are plotted as red dots and blue stars in Fig. 9(e), which shows that those of high speed images are clearly much smaller than those of the low speed images. This is caused by the motion blur of the patterns.

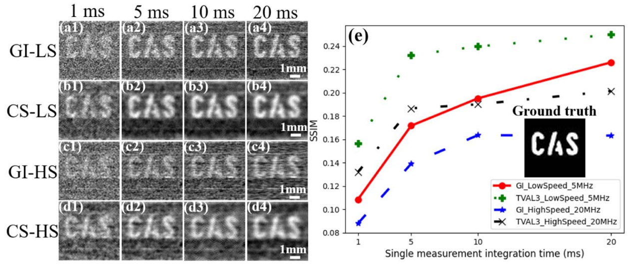

Fig. 9. (a)–(d) Images reconstructed by GI and TVAL3 algorithms at the low translation speed (GI-LS and CS-LS) of 4 mm/s and sampling frequency of 5 MHz and at the high translation speed (GI-HS and CS-HS) of 15 mm/s and sampling frequency of 20 MHz. Integration times are shown at the top of the columns. (e) SSIM curves corresponding to Figs. (a)–(d). X-ray tube current, 150 μA.

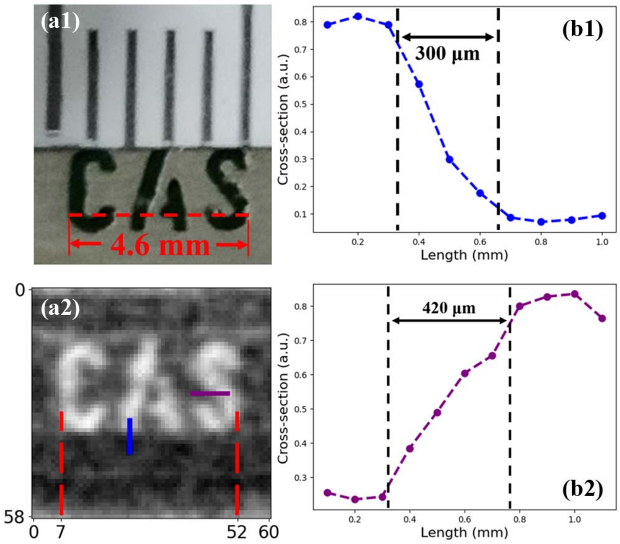

Fig. 10. (a1) Photo of the object with a ruler above showing mm divisions. (a2) Image copied from Fig. 9(b2) enlarged to the same scale as (a1). (b1) and (b2) Cross sections along the blue and purple lines of (a2).

By applying the TVAL3 algorithm, both SSIM values can be improved.

To estimate the actual achievable resolution of our setup, we compared the real size of the object [total width 4.6 mm, Fig. 10(a1)] and its corresponding width in the reconstructed image [about 46 pixels, Fig. 10(a2)]. Thus, the pixel width of the image is calculated to be ∼100 μm, which agrees well with the 100 μm width of the pattern pixels. However, the actual spatial resolution should be calculated by measuring the half-width of the slope of the cross section near the edge of the object [marked by a blue vertical line and purple horizontal line in Fig. 10(a2)] [47]. As shown in Figs. 10(b1) and 10(b2), the spatial resolutions in the horizontal and vertical directions are estimated to be approximately 150 μm and 210 μm, respectively. The former is mainly affected by the blurring effects of the uncollimated beam and other experimental issues, such as shot noise and readout noise. The latter suffers an additional motion blur effect due to the translation of the illumination patterns.

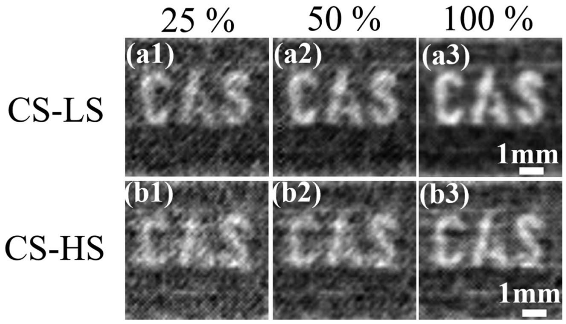

Fig. 11. Images reconstructed by TVAL3 for the (a1)–(a3) low and (b1)–(b3) high translation speeds at sampling rates of 25%, 50%, and 100%. Measurement integration time, 5 ms.

Reducing the sampling rate (the ratio of the number of sampling patterns to that of all the sampling patterns) is another method for further decreasing the exposure time. Based on CS, we obtained images for the two translation speeds at sampling rates of 25%, 50%, and 100%, for the same measurement integration time of 5 ms, as shown in Fig. 11. We can see that the images reconstructed at the low sampling rate are slightly worse in quality compared to those at high sampling rates, but the object can still be distinguished. At a 25% sampling rate, for an FOV of 5.9 mm × 6.1 mm and a speed of 15 mm/s, the exposure time is reduced from 24 s to 6 s. This proves that CS can further reduce the exposure time.

4. CONCLUSION

In contrast to existing XGI methods, we have developed a fast imaging scheme based on the concept of continuously scanning aggregate patterns. By reshaping the cyclic S-matrices to display a new measurement matrix when the scanning area moves a pixel, the imaging time can be greatly reduced. Two modulation frames of different lengths and pixel sizes were tested with different mask translation speeds, sampling frequencies, and detector integration times. The total imaging time was demonstrated to be only 4 s for an FOV of 3.4 mm × 3.8 mm and 24 s for a larger FOV of 5.9 mm × 6.1 mm. The best spatial resolution was 150 μm, which would be sufficient for preliminary clinical diagnoses. The influence of motion blur and shot noise on the reconstructed image quality is discussed. We find that the CNR and SSIM values will first increase and then decrease as the integration time increases. The more abundant the number of photons is, the shorter the optimal integration time will become. By utilizing CS, we can further improve the image quality and reduce the sampling rate to 25%. In practice, when the sampling frequency and integration time are optimized, we can also reduce the imaging time by increasing the translation speed, so long as there are sufficient photons at the detector. We have therefore demonstrated that imaging time will no longer be a problem, which resolves a critical issue for real-world applications of XGI. In addition, it is worth mentioning that the main part of our setup, the modulation mask, is an inexpensive printed circuit board, while the detector is an ordinary inexpensive X-ray integration detector.

The main contribution of this paper compared to Ref. [32] is our proposal of a fast modulation method and fabrication of the corresponding mask for X-ray ghost imaging, i.e., modulation using a physically continuously moving mask fabricated by mature PCB technology. An experimental setup is realized and its feasibility is verified. In addition, the effects of motion blur and the choice of integration time are discussed, which have not been mentioned in previous papers.

The current limitation of our setup is the mask translation speed, which may be optimized by using a rotating mask, as in Refs. [37,41]. The future limitation of our setup may be the effective photon flux. When the emitted photon flux is too high, the photon pileup effect will saturate the detector and reduce the number of effective photons detected. Thus, we need to choose a detector with a high direct current linearity upper limit [48]. Compared with our translation mask, the spinning random matrix mask in Ref. [41] has a higher modulation frequency because of its rapid rotation. However, its computing time is much longer because of the many iterations employed in CS, and its reconstructed image may be distorted due to the polar reshaping of the patterns. The rotating cyclic S-matrix mask in Ref. [37] can periodically and rapidly generate complete modulation patterns. Its short computing time and high-quality image reconstruction are more promising for future XGI setups. We believe our work represents a significant step forward toward the development of practical XGI applications.

Funding. National Natural Science Foundation of China (61975229, 12335016, 11991073, W2412039, 61805006); National Key R & D Program of China (2018YFB0504302); Strategic Priority Research Program of the CAS (XDA25030400, XDA25010100).

Disclosures. The authors declare no conflicts of interest.

Data Availability. Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

REFERENCES

1. R. Hanbury Brown, *The Intensity Interferometer: Its Application to Astronomy* (Taylor & Francis, 1974).
2. T. B. Pittman, Y. H. Shih, D. V. Strekalov, *et al.*, “Optical imaging by means of two-photon quantum entanglement,” Phys. Rev. A **52**, R3429–R3432 (1995).
3. A. V. Diebold, M. F. Imani, T. Sleasman, *et al.*, “Phaseless coherent and incoherent microwave ghost imaging with dynamic metasurface apertures,” Optica **5**, 1529–1541 (2018).
4. L. Olivieri, J. S. Totero Gongora, A. Pasquazi, *et al.*, “Time-resolved nonlinear ghost imaging,” ACS Photonics **5**, 3379–3388 (2018).
5. C. M. Watts, D. Shrekenhamer, J. Montoya, *et al.*, “Terahertz compressive imaging with metamaterial spatial light modulators,” Nat. Photonics **8**, 605–609 (2014).
6. L. Olivieri, J. S. T. Gongora, L. Peters, *et al.*, “Hyperspectral terahertz microscopy via nonlinear ghost imaging,” Optica **7**, 186–191 (2020).
7. V. Kumar, V. Cecconi, L. Peters, *et al.*, “Deterministic terahertz wave control in scattering media,” ACS Photonics **9**, 2634–2642 (2022).
8. V. Cecconi, V. Kumar, J. Bertolotti, *et al.*, “Terahertz spatiotemporal wave synthesis in random systems,” ACS Photonics **11**, 362–368 (2024).
9. V. Kumar, V. Cecconi, A. Cutrona, *et al.*, “Terahertz microscopy through complex media,” arXiv, arXiv:2412.06427 (2024).
10. D. Pelliccia, A. Rack, M. Scheel, *et al.*, “Experimental X-ray ghost imaging,” Phys. Rev. Lett. **117**, 113902 (2016).
11. H. Yu, R. Lu, S. Han, *et al.*, “Fourier-transform ghost imaging with hard X rays,” Phys. Rev. Lett. **117**, 113901 (2016).
12. A.-X. Zhang, Y.-H. He, L.-A. Wu, *et al.*, “Tabletop X-ray ghost imaging with ultra-low radiation,” Optica **5**, 374–377 (2018).
13. Y.-H. He, A.-X. Zhang, M.-F. Li, *et al.*, “High-resolution sub-sampling incoherent X-ray imaging with a single-pixel detector,” APL Photonics **5**, 056102 (2020).
14. R. I. Khakimov, B. M. Henson, D. K. Shin, *et al.*, “Ghost imaging with atoms,” Nature **540**, 100–103 (2016).
15. S. Li, F. Cropp, K. Kabra, *et al.*, “Electron ghost imaging,” Phys. Rev. Lett. **121**, 114801 (2018).
16. A. M. Kingston, G. R. Myers, D. Pelliccia, *et al.*, “Neutron ghost imaging,” Phys. Rev. A **101**, 053844 (2020).
17. Y.-H. He, Y.-Y. Huang, Z.-R. Zeng, *et al.*, “Single-pixel imaging with neutrons,” Sci. Bull. **66**, 133–138 (2021).
18. J.-T. Xie, J.-H. Tan, S.-L. Wang, *et al.*, “Simultaneous imaging and element differentiation by resonant neutron ghost imaging,” Front. Phys. **20**, 32203 (2025).
19. T. J. Lane and D. Ratner, “What are the advantages of ghost imaging? Multiplexing for X-ray and electron imaging,” Opt. Express **28**, 5898–5918 (2020).
20. J. Cheng, “Ghost imaging through turbulent atmosphere,” Opt. Express **17**, 7916–7921 (2009).
21. J.-E. Oh, Y.-W. Cho, G. Scarcelli, *et al.*, “Sub-rayleigh imaging via speckle illumination,” Opt. Lett. **38**, 682–684 (2013).
22. J. H. Shapiro, “Computational ghost imaging,” Phys. Rev. A **78**, 061802 (2008).
23. M. F. Duarte, M. A. Davenport, D. Takhar, *et al.*, “Single-pixel imaging via compressive sampling,” IEEE Signal Process. Mag. **25**, 83–91 (2008).
24. A. Schori and S. Shwartz, “X-ray ghost imaging with a laboratory source,” Opt. Express **25**, 14822–14828 (2017).
25. D. Pelliccia, M. P. Olbinado, A. Rack, *et al.*, “Towards a practical implementation of X-ray ghost imaging with synchrotron light,” IUCrJ **5**, 428–438 (2018).
26. A. M. Kingston, D. Pelliccia, A. Rack, *et al.*, “Ghost tomography,” Optica **5**, 1516–1520 (2018).
27. K. Tamasaku, T. Sato, T. Osaka, *et al.*, “Dynamically patterning X-ray beam by a femtosecond optical laser,” Sci. Adv. **10**, eadp5326 (2024).
28. D. Ceddia and D. M. Paganin, “Random-matrix bases, ghost imaging, and X-ray phase contrast computational ghost imaging,” Phys. Rev. A **97**, 062119 (2018).
29. M. P. Olbinado, D. M. Paganin, Y. Cheng, *et al.*, “X-ray phase contrast ghost imaging using a single-pixel camera,” Optica **8**, 1538–1544 (2021).
30. Y. Klein, O. Sefi, H. Schwartz, *et al.*, “Chemical element mapping by X-ray computational ghost fluorescence,” Optica **9**, 63–70 (2022).
31. J.-T. Xie, J.-H. Tan, S.-H. Bie, *et al.*, “Simultaneous imaging and element differentiation by energy-resolved X-ray absorption ghost imaging,” Opt. Lett. **49**, 4162–4165 (2024).
32. P. Kilcullen, T. Ozaki, and J. Liang, “Compressed ultrahigh-speed single-pixel imaging by swept aggregate patterns,” Nat. Commun. **13**, 7879 (2022).
33. W. Huda and R. B. Abrahams, “X-ray-based medical imaging and resolution,” Am J. Roentgenol. **204**, W393–W397 (2015).
34. M.-J. Sun, W. Chen, T.-F. Liu, *et al.*, “Image retrieval in spatial and temporal domains with a quadrant detector,” IEEE Photonics J. **9**, 3901206 (2017).
35. B. I. Erkmen and J. H. Shapiro, “Ghost imaging: from quantum to classical to computational,” Adv. Opt. Photonics **2**, 405–450 (2010).
36. M. Harwit and N. J. A. Sloane, *Hadamard Transform Optics* (Elsevier, 1979).
37. E. Hahamovich, S. Monin, Y. Hazan, *et al.*, “Single pixel imaging at megahertz switching rates via cyclic Hadamard masks,” Nat. Commun. **12**, 4516 (2021).

38. A. M. Kingston, A. Aminzadeh, L. Roberts, *et al.*, "Optimizing nonconfigurable, transversely displaced masks for illumination patterns in classical ghost imaging," Phys. Rev. A **107**, 023524 (2023).
39. A. Aminzadeh, L. Roberts, B. Young, *et al.*, "Mask design, fabrication, and experimental ghost imaging applications for patterned X-ray illumination," Opt. Express **31**, 24328–24346 (2023).
40. W. Jiang, X. Li, X. Peng, *et al.*, "Imaging high-speed moving targets with a single-pixel detector," Opt. Express **28**, 7889–7897 (2020).
41. W. Jiang, J. Jiao, Y. Guo, *et al.*, "Single-pixel camera based on a spinning mask," Opt. Lett. **46**, 4859–4862 (2021).
42. E. J. Candes and J. K. Romberg, "Signal recovery from random projections," Proc. SPIE **5674**, 76–86 (2005).
43. D. Donoho, "Compressed sensing," IEEE Trans. Inf. Theory **52**, 1289–1306 (2006).
44. C. Li, W. Yin, H. Jiang, *et al.*, "An efficient augmented Lagrangian method with applications to total variation minimization," Comput. Optim. Appl. **56**, 507–530 (2013).
45. O. Katz, Y. Bromberg, and Y. Silberberg, "Compressive ghost imaging," Appl. Phys. Lett. **95**, 131110 (2009).
46. P. Zerom, Z. Shi, M. N. O'Sullivan, *et al.*, "Thermal ghost imaging with averaged speckle patterns," Phys. Rev. A **86**, 063817 (2012).
47. O. Sefi, A. B. Yehuda, Y. Klein, *et al.*, "20 µm resolution multipixel ghost imaging with high-energy X-rays," Opt. Express **32**, 37001–37010 (2024).
48. Hamamatsu, *Photomultiplier Tubes Basics and Applications*, 4th ed. (Hamamatsu Photonics, 2017).