

High-resolution sub-sampling incoherent x-ray imaging with a single-pixel detector

Cite as: APL Photonics 5, 056102 (2020); <https://doi.org/10.1063/1.5140322>

Submitted: 26 November 2019 . Accepted: 13 April 2020 . Published Online: 04 May 2020

Yu-Hang He , Ai-Xin Zhang, Ming-Fei Li, Yi-Yi Huang, Bao-Gang Quan, Da-Zhang Li, Ling-An Wu , and Li-Ming Chen 

COLLECTIONS

Paper published as part of the special topic on [Photonics and AI](#)

Note: This article is part of the APL Photonics Special Topic on Photonics and AI.



View Online



Export Citation



CrossMark

ARTICLES YOU MAY BE INTERESTED IN

[Virtual optofluidic time-stretch quantitative phase imaging](#)

APL Photonics 5, 046103 (2020); <https://doi.org/10.1063/1.5134125>

[InAs-based quantum cascade lasers grown on on-axis \(001\) silicon substrate](#)

APL Photonics 5, 041302 (2020); <https://doi.org/10.1063/5.0002376>

[Ghost imaging second harmonic generation microscopy](#)

Applied Physics Letters 116, 191101 (2020); <https://doi.org/10.1063/1.5144690>

APL Photonics The Future Luminary Award

Journal
Impact Factor
4.383

LEARN MORE!

High-resolution sub-sampling incoherent x-ray imaging with a single-pixel detector

Cite as: APL Photon. 5, 056102 (2020); doi: 10.1063/1.5140322

Submitted: 26 November 2019 • Accepted: 13 April 2020 •

Published Online: 4 May 2020



View Online



Export Citation



CrossMark

Yu-Hang He,^{1,2}  Ai-Xin Zhang,^{1,2} Ming-Fei Li,^{3,4} Yi-Yi Huang,^{1,2} Bao-Gang Quan,^{1,2} Da-Zhang Li,⁵ Ling-An Wu,^{1,2,a)}  and Li-Ming Chen^{2,6,a)} 

AFFILIATIONS

¹Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China

²University of Chinese Academy of Sciences, Beijing 100049, China

³Beijing Institute of Aerospace Control Devices, Beijing 100039, China

⁴Quantum Engineering Research Center, China Aerospace Science and Technology Corporation, Beijing 100094, China

⁵Institute of High Energy Physics, Chinese Academy of Sciences, Beijing 100049, China

⁶Department of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai 200240, China

Note: This article is part of the APL Photonics Special Topic on Photonics and AI.

a) Authors to whom correspondence should be addressed: wula@iphy.ac.cn and lmchen@sjtu.edu.cn

ABSTRACT

X-ray “ghost” imaging has drawn great attention for its potential to obtain images with a high resolution and lower radiation dose in medical diagnosis, even with only a single-pixel detector. However, it is hard to realize with a portable x-ray source due to its low flux. Here, we demonstrate a computational x-ray ghost imaging scheme where a real bucket detector and specially designed high-efficiency modulation masks are used, together with a robust deep learning algorithm in which a compressed set of Hadamard matrices is incorporated into a multi-level wavelet convolutional neural network. With a portable incoherent x-ray source of $\sim 37 \mu\text{m}$ diameter, we have obtained an image of a real object from only 18.75% of the Nyquist sampling rate. A high imaging resolution of $\sim 10 \mu\text{m}$ has been achieved, which is required for cancer detection and so represents a concrete step toward the realization of a practical low cost x-ray ghost imaging camera for applications in biomedicine, archeology, material science, and so forth.

© 2020 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/1.5140322>

I. INTRODUCTION

Safety and image quality are the two major factors in x-ray imaging. In traditional schemes, according to the Ross criterion,¹ there has to be a balance between radiation dose and image quality since high resolution and good contrast require a sufficiently long exposure, which means greater dose. Brilliant phase contrast images with nanometer resolution can be obtained in state-of-the-art synchrotron facilities that provide monochromatic ultra-bright x-ray beams.^{2,3} For most users, however, incoherent x-ray sources for use in a laboratory are more accessible, but compared with synchrotron or free-electron laser sources, they are much weaker, and conventional focusing or diffractive optical elements are not applicable. As a result, straightforward shadow projection microscopes

are widely used, especially in clinical medical imaging, in which case the resolution is mainly limited by the source size.^{4,5} In raster scanning transmission x-ray microscopy, high resolution images can be obtained by focusing a soft x-ray beam onto a spot, which is then scanned over the whole sample. However, focusing is complicated, costly, and greatly attenuates the beam, again necessitating a high intensity source. Hence, how to increase the resolution while lowering the cost of x-ray imaging with such polychromatic sources is a significant problem.

Different from traditional imaging, ghost imaging (GI) is a second-order correlation based technology that retrieves information about an object from a series of reference patterns and the corresponding intensity values measured by using a single-pixel (“bucket”) detector.⁶ As this type of detector is more sensitive than

an array of pixels, even in poor weather⁷ or ultra-low exposure situations,⁸ GI can still retrieve an acceptable image. The key point is to generate or record a series of reference speckle patterns that illuminate the object plane; then by convoluting the total intensity measured at the bucket detector with the reference patterns, the image can be retrieved. Since the speckle patterns are similar to the matrix arrays in pixel detectors, the resolution is limited by the average size of the speckles. However, through various means, GI can now achieve resolutions beyond the Rayleigh criterion⁹ and has already been applied to many fields such as microscopy,^{10,11} lidar detection, and remote sensing.^{12–14}

Over the past few decades, GI has been demonstrated with quantum light,⁶ classical light,¹⁵ and even particles—atoms¹⁶ and electrons.¹⁷ There is great potential in the fabrication of cheap high resolution GI cameras at terahertz¹⁸ and infrared¹⁰ wavelengths that cannot use silicon sensors. Ghost tomography has broad prospects in various fields.^{19–21} At visible wavelengths, it is easy to obtain reference patterns with the aid of a beamsplitter, which can further be replaced by a spatial light modulator and just the bucket detector in a single light path, as in computational GI.^{22,23} At x-ray wavelengths, however, a major difficulty is that there are no suitable optics. Several approaches have been proposed to overcome this problem. In the beam splitting strategy,^{24–27} a

crystal can be taken as a beamsplitter based on Laue diffraction if the flux density of the source in a narrow bandwidth is very strong; even so, mechanical vibration may blur the image if not controlled strictly. Another strategy is pre-recording,^{28,29} with a phase or amplitude modulation plate, a series of repeatable speckle patterns can be generated quite easily, but saving and transferring the large amount of data required entails much extra time, which greatly reduces practicability. Another problem is that all the existing schemes rely on a high resolution x-ray camera for calibration, which increases the cost, while the resolution is limited by the pixel size of the CCD (or CMOS) arrays. Recently, x-ray GI of a one-dimensional slit has been realized with a single-pixel detector, but with the bright monochromatic beam from a synchrotron source.³⁰ For common practical applications such as in biomedicine, a resolution of several micrometers must be achieved to be of real use in diagnosis.

II. RESULTS

In this letter, we report a computational x-ray GI (CXGI) scheme by which high resolution images were obtained with only an inexpensive single-pixel detector and a portable incoherent low flux x-ray source, plus the use of a deep learning algorithm. Instead of our

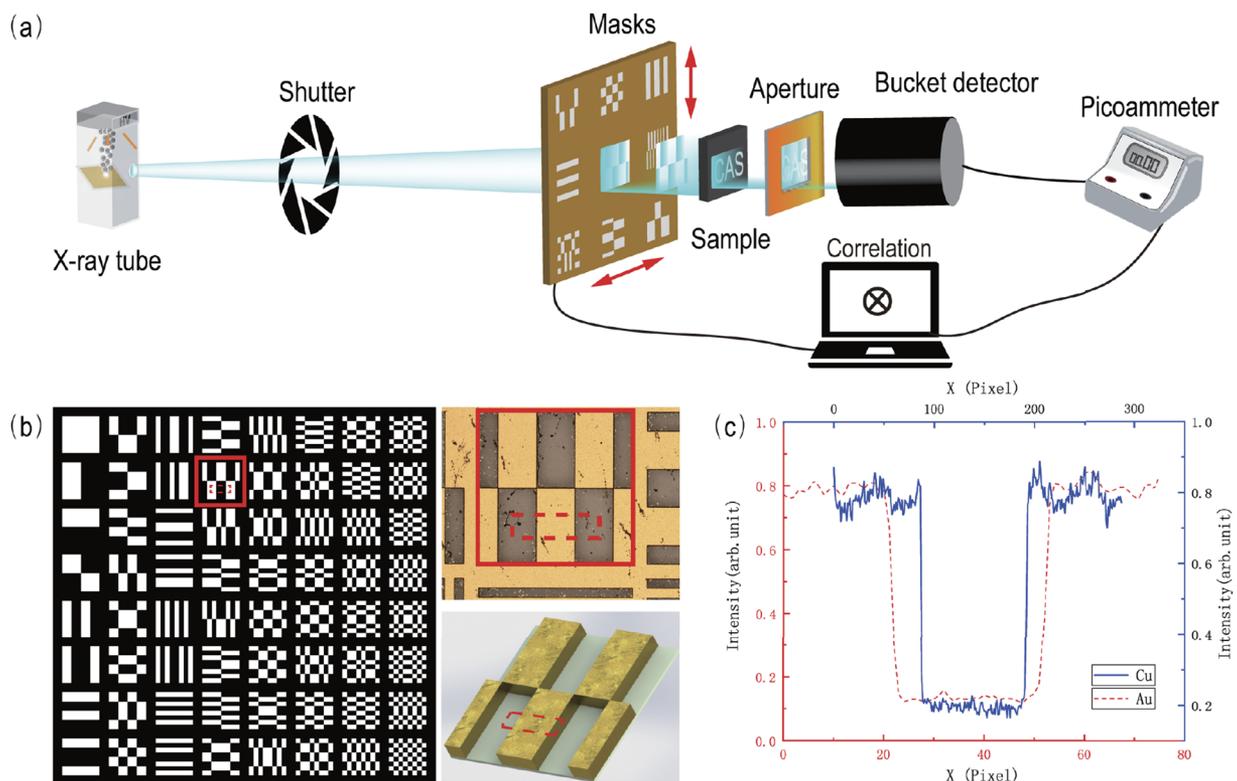


FIG. 1. (a) Experimental setup of CXGI. The set of masks is mounted on a motorized 2D translation stage. (b) A typical set of the specially designed Hadamard matrices; one of the patterns (enclosed in red) for the Au mask is shown enlarged in the upper right, and its 3D visualization in the lower right. (c) Cross section of the portion enclosed by the red dashed line in (b); blue solid line: Cu on a laminate substrate with a modulation depth ratio of 75% and red dashed line: Au on a SiO₂ substrate with a modulation ratio of 83%.

previous scheme in which randomly modulated patterns were pre-recorded by using an array detector,²⁹ a transmission mask engraved with custom designed orthogonal patterns was fabricated to provide amplitude modulation. The layout of the experiment is shown in Fig. 1(a). As shown, the incoherent hard x-ray beam from a micro-focus x-ray tube (Incoatec Source I_{μs}) first passes through an adjustable shutter and then through a certain matrix in the modulation mask, which is mounted on a two-dimensional motor stage. The modulated x-ray pattern illuminates the sample, which is partially transmitting, and its intensity is measured by using a bucket detector. Another aperture blocks out unwanted light from around the object and helps to identify the field of view. Since only the total intensity has to be recorded, the distance between the object and the detector is quite flexible and can be made as far apart as convenient, which will further reduce the influence of undesirable scattered photons. After each exposure, the shutter is closed, the mask is translated to the next adjacent matrix, and the measurement is repeated. The single-pixel detector used in our experiment was an x-ray diode (Hamamatsu) with a beryllium window and wrapped in a copper shell, which converted the incident intensity to a current of several pico-amperes. Of course, a CCD camera could be used as a bucket detector by integrating the total intensity registered on all the pixels, but this requires a much longer processing time and is much more expensive, while a true bucket detector can have better sensitivity and is more efficient for data collection and processing. After several measurements, the image is retrieved by an appropriate algorithm.

The amplitude modulation board was fabricated from a metal layer etched into a series of patterns upon a flat substrate, the former being strongly absorbing and the latter being transparent to hard

x rays. Two different masks were made: the first was an inexpensive printed circuit board composed of a 100 μm thick copper foil on a 500 μm thick laminate substrate; this was used to test the feasibility of our scheme. The second consisted of a 10 μm thick layer of gold foil electroplated onto a 4 in. square, 500 μm thick SiO₂ substrate; this board was made for high resolution CXGI. Both masks were etched with a set of Hadamard matrices;³¹ the number and size of the pixels of the Cu and Au masks were 32 × 32, 150 μm and 64 × 64, 10 μm, respectively. A detailed description of the matrix design is given in the [supplementary material](#), Sec. 3. An illustration of part of the masks is shown in Fig. 1(b). An enlargement of the area outlined by the red solid line for the Au mask is shown in the upper right, and the lower right is the corresponding three-dimensional (3D) visualization. We define the modulation depth ratio D_r as $D_r = (D_{max} - D_{min})/D_{max}$, where D_{max} and D_{min} represent the maximum and minimum intensities recorded by using an x-ray CCD (Andor iKon-M), respectively. The cross section of the part enclosed by the red dashed line in Fig. 1(b) is shown in Fig. 1(c), from which we can see that the modulation depth ratio D_r of the Cu and Au masks is about 75% and 83%, respectively. This profile was plotted from the gray values of a direct image on a CCD camera of the x-ray transmission through the mask.

When the Cu mask was used, a 5 mm thick stainless steel stencil with the letters “CAS” cut out was chosen as the object, as shown in Fig. 2(a). We compared the performance of two different modulation means: random sandpaper speckles and a set of pre-designed Hadamard matrices. The same second-order correlation algorithm was taken for fairness. By adjusting the magnification and speckle size, the resolution was set at 150 μm for both cases. We adopt the contrast-to-noise ratio (CNR) as a criterion of image quality, defined

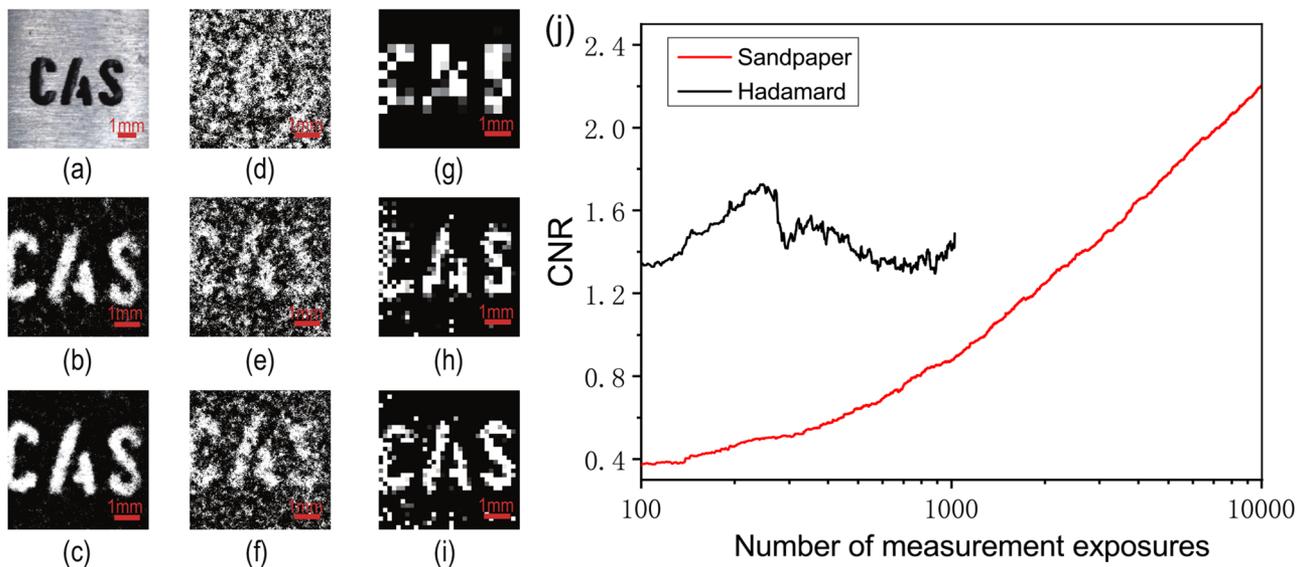


FIG. 2. (a) Object: 5 mm thick stainless steel stencil with the letters “CAS” cut out. [(b)–(f)] Sandpaper x-ray GI images retrieved after 5000, 10 000, 128, 512, and 1024 exposures. [(g)–(i)] CXGI images retrieved with the Cu mask after 128, 512, and 1024 exposures. (j) CNR vs number of exposures for x-ray GI (red line) and CXGI (black line).

as

$$CNR \equiv \frac{\langle G_1 \rangle - \langle G_0 \rangle}{\sqrt{\sigma_1^2 + \sigma_0^2}}, \quad (1)$$

where G_1 and G_0 are the GI values for any pixel where the transmission is 1 or 0, respectively; σ_1^2 and σ_0^2 are the corresponding variances, i.e., $\sigma_1^2 = \langle G_1^2 \rangle - \langle G_1 \rangle^2$.

When sandpaper is used, since the modulation is random and its speckles are quite uniformly distributed, it can perform well when the number of exposures equals or exceeds the number of pixels. This can be seen from Figs. 2(b) and 2(c), which correspond to 5000 and 10000 exposures, respectively; the retrieved image contains about 3500 pixels in this case. Figures 2(d)–2(f) and Figs. 2(g)–2(i) show the results of the sandpaper speckles and Hadamard masks for 128, 512, and 1024 measurements, respectively. We see that the Hadamard mask performs much better than the sandpaper under the same number of measurements. Figure 2(j) provides a more quantitative comparison, where the CNR of both methods is plotted as a function of the number of exposures. Here, for the same CNR, when the Hadamard board is used, the number of measurements is reduced by an order of magnitude due to the orthogonality of the mask patterns. However, the CNR begins to decrease when the exposure number exceeds 300. In addition, it is evident from Fig. 2(i) that even in the full Nyquist sampling case, there is still some unwanted noise. There are two explanations for this: one is imperfections of the Hadamard patterns due to uneven etching during the electrochemical processing, which become more pronounced in the

finer patterns containing more complex structures; the other is that the finer structures produce smaller fluctuations of the x-ray intensity, which therefore cannot be easily detected when the detector is not sensitive enough. The first problem could be solved with more precise lithography, and the second by either increasing the source intensity or improving the sensitivity of the detector, e.g., by using a photomultiplier tube.

Aside from the updates in the hardware device, the CXGI image can also be further improved by specially designed algorithms. Compressed sensing (CS) has been widely used in GI and single-pixel cameras.^{32–34} For the same bucket signals of the “CAS” object mentioned above, Fig. 3(c) shows the image reconstructed by the total variation augmented Lagrangian alternating direction algorithm (TVAL3), one of the most popular CS algorithms.³⁵ The image is greatly improved as it has less noise, and the edges are much sharper than in Fig. 3(b). The CNR values of Figs. 3(b) and 3(c) are 1.49 and 1.75, respectively. To test the universality of this algorithm, we take a more complex sample as the object, which is a metal gear with 14 teeth, about 1.5 mm in diameter. The CNR of Fig. 3(g) retrieved by TVAL3 is 0.5, which is even a little bit poorer than 0.6 of Fig. 3(f) obtained by traditional GI. It seems that the CS algorithm is not robust enough when the image contains a complex structure and the Hadamard mask is imperfect.

Deep neural networks are computational models that learn representations of data with multiple levels of abstraction.³⁶ They are proven to be very successful at discovering features in high-dimensional data in many areas, including GI. Recently, deep

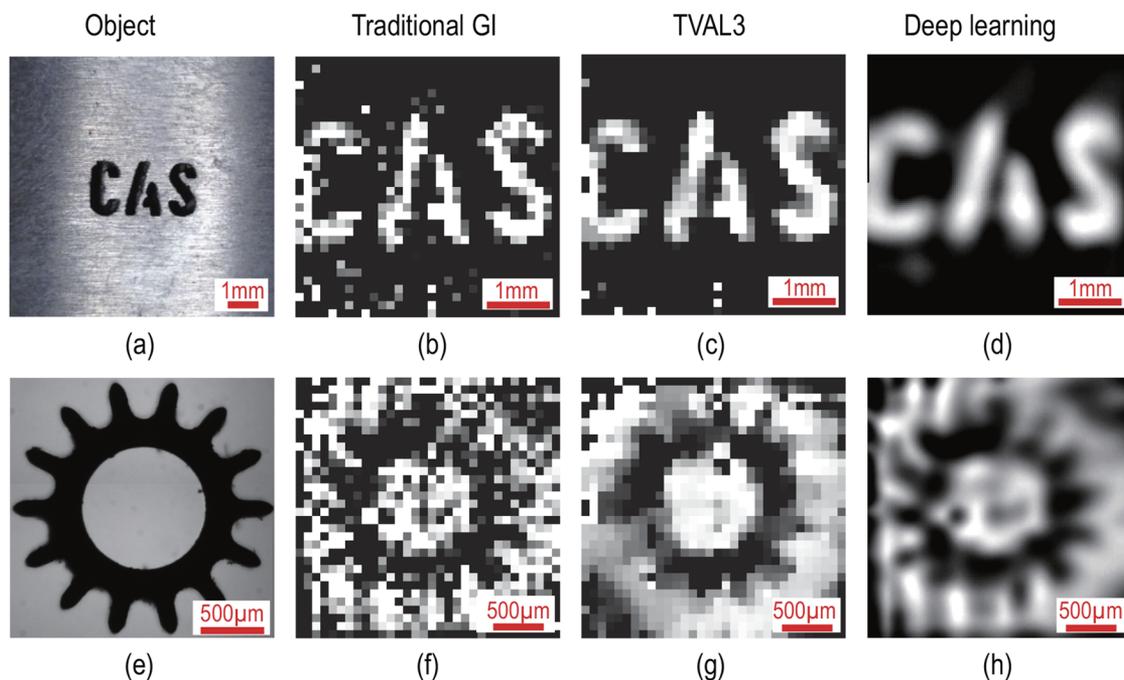


FIG. 3. Results of x-ray GI with the Cu modulation mask and their corresponding images. (a) Same object as in Fig. 2(a). [(b)–(d)] Images retrieved by traditional GI, TVAL3, and deep learning after 1024 exposures, with CNR values of 1.49, 1.75, and 2.43, respectively. (e) Object: a metal gear with 14 teeth. [(f)–(h)] Images of metal gear retrieved by traditional GI, TVAL3, and deep learning after 1024 exposures, with CNR values of 0.6, 0.5, and 1.56, respectively.

learning was successfully used to recover structured signals (in particular, images) from their under-sampled random linear measurements.^{37–39} With an irregular trained basis, it can process an image at high speed even under a 2% exposure compression ratio.⁴⁰ However, the rectangular Hadamard matrices that we chose (because these shapes are relatively easy to fabricate) are not suitable for the deep convolutional auto-encoder network used in the deep learning algorithm mentioned above. Thus, we developed a new compressible Hadamard plus multi-level wavelet convolutional neural network (CH-MWCNN) algorithm, which takes the receptive field size and computational efficiency into consideration and can be rather robust even with an imperfect modulation mask, as in our case. With this, we succeeded in obtaining greatly improved images, as shown in Figs. 3(d) and 3(h), where the CNR ratios are 2.43 and 1.56, respectively, higher than all the other methods. Admittedly, there is still some blurring present due to the fact that the size of the modulation mask is only 32×32 pixels, which means that the sparser pixels cannot bring out enough detail, while the finer structures tend to be over-fitted by our CH-MWCNN algorithm. This could be greatly improved by increasing the density of the pixels although, of course, this would require better precision in the mask fabrication. Further details of the effect of noise on the algorithm are given in the [supplementary material](#).

To improve both hardware and software, the Cu mask was replaced by a 64×64 Au mask with pixels of size $10 \mu\text{m}^2$. A

semi-cylindrically shaped object made of gold and glued onto a rectangular column was used here as our object. Its 3D visualizations from different angles are shown in Fig. 4(a); the exposed area of the object was actually just $0.64 \times 0.64 \text{ mm}^2$. A direct absorption/projection x-ray image of the object is presented in Fig. 4(c), where the exposure time was 5 s. Although the gap is visible, there are many noisy dots widely distributed throughout the whole image, which makes it hard to resolve the true distance between the semi-cylinder and rectangular column. Figure 4(d) shows the image recovered by TVAL3; we can see that it is seriously blurred, and the CNR is only 0.27. Similar to the problems in the manufacturing of the Cu mask, there is also distortion in the ion etched Au mask, part of which is shown in the scanning electron microscope (SEM) image of Fig. 4(b) taken at an angle of 52° . Here, we observe the sloping profile of the etched edges, which ideally should be perpendicular. This is probably the reason why the TVAL3 algorithm failed to give better detail. On the other hand, when CH-MWCNN was used, the CNR of the retrieved image, shown in Fig. 4(e), improved significantly; here, it is 2.65 and there are many fewer noisy dots. The gap between the semi-cylinder and the rectangular column, which is $\sim 10 \mu\text{m}$ wide, can be distinguished clearly [note that Figs. 4(c)–4(e) show 2D transmission images of the 3D object]. This result was obtained under a sampling rate of 18.75% with 0.3 s recovery time, a performance much better than TVAL3, and certain other compressed sensing algorithms. The average processing time of

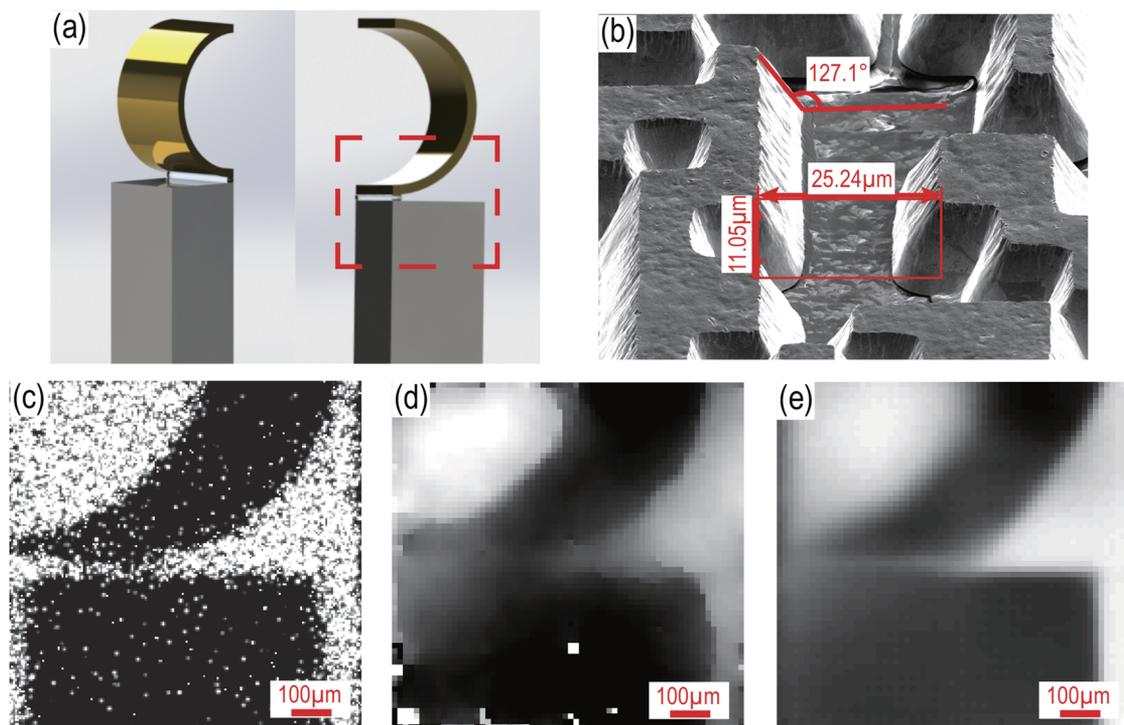


FIG. 4. Object and images with a gold modulation mask. (a) Two 3D views of the object from different angles where the area enclosed by the red dashed line is the actually exposed part of the object. (b) SEM image of part of the Au mask. (c) Direct x-ray absorption/projection image of the sample with 5 s exposure time. (d) CXGI image retrieved by TVAL3 after 4096 exposures; its CNR is 0.27. (e) Image reconstructed by CH-MWCNN after 768 exposures; its CNR is 2.65.

CH-MWCNN was between 0.2 s and 1 s for an image of 64×64 pixels, depending on its complexity, running on a laptop with an Intel® Core™ i7-6600U central processing unit and 12 GB random access memory.

III. DISCUSSION

In shadow projection imaging, how to ensure both magnification and resolution is a difficult problem because according to geometrical optics the object's penumbra will blur the image when the object is smaller than half the source size; this is the key limit of resolution in both traditional absorption and propagation based phase contrast x-ray imaging. The x-ray source size in our experiment was $30 \times 37 \mu\text{m}^2$, measured by a knife edge method, so it would be difficult to distinguish details as fine as $19 \mu\text{m}$ directly. A possible solution would be to decrease the source size by a pin-hole, but this would reduce the flux significantly, or to perform focused raster scanning, but this requires a monochromatic beam as well as sophisticated optics. In contrast, the Hadamard patterns in our CXGI scheme have high transmittance, while a bucket detector can be much more sensitive than the array detectors used in traditional x-ray microscopy; thus, our scheme is very suitable for a low-flux source. The image quality of CXGI depends on the design and quality of the illumination masks, which in our case are quite good. The resolution of our current experiments is, in fact, limited by the mask lithography technology, and so far we have achieved a value of several micrometers, as shown in Fig. 4. Our low sampling rate means lower dose, less measurement time, and faster processing, which are all essential for real applications. Of course the image quality needs to be improved further, but the real objects imaged by our CH-MWCNN algorithm fully indicate the huge potential of our CXGI scheme. If applied to a certain field such as medical diagnosis where training can be acquired with the vast clinical image data resources available, the results should certainly be much better.

In conclusion, we have realized CXGI with an incoherent x-ray source and a true bucket detector, with both simulation and experimental results showing that we have surpassed the resolution limit of incoherent x-ray imaging even at subsampling rates. The setup is simple, cost effective, and convenient to operate. Compared with random speckle modulation, the pre-designed masks have a consistent speckle size so that the resolution can be predetermined, and combined with certain orthogonal matrices such as Hadamard matrices, the measurement efficiency can be greatly improved. A new CH-MWCNN algorithm has been implemented, by which even when the modulation mask contains some distortions, we can still observe fine structures under a low subsampling rate, regardless of the complexity of the object. As a result, images with $10 \mu\text{m}$ resolution have been obtained for a source size of about $37 \mu\text{m}$ at a sampling rate of 18.75%, which indicates that both the measurement time and the radiation dosage in x-ray diagnosis can be greatly reduced. The resolution could be improved further if finer and better masks were used, in addition to optimization of other technical aspects. With more sensitive photodiodes and less noisy electronics, we should be able to reduce the dosage down to single-photon levels. The image quality would also be much better if imaging data for real samples could be used to strengthen our deep learning algorithm. Already, our current CXGI scheme demonstrates

that it should be quite feasible to build a practical, low-cost, single-pixel x-ray camera for use in biomedicine, archeology, and material diagnosis.

SUPPLEMENTARY MATERIAL

See the [supplementary material](#) for more details.

AUTHOR'S CONTRIBUTIONS

Y.-H.H. and A.-X.Z. contributed equally to this work.

ACKNOWLEDGMENTS

We thank Professor Junjie Li for technical support during fabrication of the modulation masks. This work was supported by the National Key R&D Program of China (Grant Nos. 2017YFA0403301, 2017YFB0503301, and 2018YFB0504302), the National Natural Science Foundation of China (Grant Nos. 11721404, 11991073, 61975229, 61805006, and 11805266), the Key Program of CAS (Grant Nos. XDB17030500 and XDB16010200), and the Science Challenge Project (No. TZ2018005).

APPENDIX: METHODS

1. Experimental setup

A schematic diagram of the experimental setup is presented in Fig. 1(a). The hard x-ray source is a portable x-ray tube (Incoatec Microfocus Source I μ s); when operating at 40 kV and 600 μA , it emits polychromatic x rays composed mainly of K α radiation with a characteristic central wavelength of 0.15 nm (corresponding to 8.04 keV photon energy). The object was placed at a distance of 45.5 cm from the source, and a detector was placed 5 cm behind the object. For a fair comparison, these distances were kept the same in both the direct projection and CXGI experiments although an x-ray CCD was used for the former and a single-pixel detector setup for the latter. The modulation patterns projected onto the object are shadows of pre-designed masks, created by absorption and transmission, not interference. Since the beam size and its divergence angle (0.1°) are rather small, the masks had to be sufficiently far from the source so that the beam could cover an entire sample; the distances were 168 cm and 44.5 cm for the Cu and Au masks, respectively, with the sample placed 1 cm behind, as close as convenient, to avoid magnification and ensure resolution. Since there has to be a finite horizontal and vertical spacing between each matrix, three pixels for the copper mask and six pixels for the gold mask, this would transmit spurious random x rays into the bucket detector. A square aperture was therefore put behind the sample to block unwanted signals. The modulation board was mounted on a two-dimensional motorized translation stage that could be controlled to a precision of $0.1 \mu\text{m}$. A different Hadamard mask was used for each illumination of the sample, so the stage had to be translated through $(4.8 + 0.45) = 5.25 \text{ mm}$ (Cu mask case) or $(0.64 + 0.06) = 0.7 \text{ mm}$ (Au mask case) after each measurement; this was performed by using a computer program. Every bucket signal was converted to a current signal that was read out by using a Keithley Model 6485 picoammeter, which was set in the low speed mode to reduce the readout noise since the

response time is about 300 ms. As our x-ray source is continuous, a shutter was used to block the beam and opened only during image exposure.

2. The CH-MWCNN algorithm

The CH-MWCNN algorithm consists of two parts, the encoding and decoding layers. The main function of the former is to sort out the best arrangement of the masks according to their sparsity based on a large amount of image datasets so that although the less sparse masks contain more distortion, they can still produce a high quality image by maintaining the balance between the measurements and the noise arising from processing distortion. In this layer, 90 000 images were used to train the arrangement of the Hadamard bases, each image being convoluted with the whole set of matrices. For each resulting image, its signal is expressed as a 4096×1 vector y_i , where $i = 1$ to 90 000. The Hadamard bases are then rearranged in descending order according to the absolute value of the summed signals $\beta = |y_i|$. The first 768 rearranged Hadamard bases indexed by 768 elements of β (those with the highest value) are retained. In our experiment, the selected 768 Hadamard bases are used to measure the real sample, and a vector $y_{768 \times 1}$ representing all the bucket signals is obtained. In the decoding layer, this $y_{768 \times 1}$ vector is padded to $y_{4096 \times 1}$ with zeros and indexed in the Walsh order; then, a fast Walsh–Hadamard transform is used to recover the image vector, and a final 64×64 image acquired after reshaping and normalizing. From this preliminary reconstructed 64×64 image, the multi-level wavelet convolutional neural network⁴⁰ is used to further improve the reconstruction of the image, and the output is our final image.

3. Micro-manufacture of modulation masks

The patterns in our Cu mask were fabricated on a 3 oz printed circuit board by wet etching, a common industrial process. However, creation of the finer $10 \mu\text{m}$ resolution Au mask was much more complex and required photolithography combined with ion beam etching. A 4 in. diameter quartz wafer was coated by electron evaporation with a seed layer of 10 nm titanium and 60 nm gold, followed by electrochemical deposition of a $10 \mu\text{m}$ thick film of gold. Then, the wafer was coated with AZ4620 photoresist and soft-baked at 100°C for 3 min for solvent removal prior to patterning. The photoresist was, in turn, patterned using a Karl Suss MA6 Contact Aligner and photomask with 250 mW/cm^2 of uv exposure at 365 nm. The patterned wafer was then immersed in the developer (AZ400k: deionized water = 1:3) for 2 min. Finally, the photoresist pattern was transferred onto the gold film via ion beam etching by using an Ar ion milling system (LKJ-150, Beijing Institute of Advanced Ion Beam Technology) with an ion energy of 300 eV and ion current of 0.5 mA/cm^2 .

REFERENCES

- ¹A. Rose, "The visual process," in *Vision: Optical Physics and Engineering* (Springer, Boston, MA, 1973).
- ²U. Bonse and M. Hart, *Appl. Phys. Lett.* **6**, 155 (1965).
- ³H. Wen, A. A. Gomella, A. Patel, S. K. Lynch, N. Y. Morgan, S. A. Anderson, E. E. Bennett, X. Xiao, C. Liu, and D. E. Wolfe, *Nat. Commun.* **4**, 2659 (2013).
- ⁴A. Pogany, D. Gao, and S. W. Wilkins, *Rev. Sci. Instrum.* **68**, 2774 (1997).

- ⁵P. Cloetens, R. Barrett, J. Baruchel, J.-P. Guigay, and M. Schlenker, *J. Phys. D: Appl. Phys.* **29**, 133 (1996).
- ⁶T. B. Pittman, Y. H. Shih, D. V. Strelakov, and A. V. Sergienko, *Phys. Rev. A* **52**, R3429 (1995).
- ⁷R. E. Meyers, K. S. Deacon, and Y. Shih, *Appl. Phys. Lett.* **98**, 111115 (2011).
- ⁸G. A. Howland, D. J. Lum, M. R. Ware, and J. C. Howell, *Opt. Express* **21**, 23822 (2013).
- ⁹X.-H. Chen, F.-H. Kong, Q. Fu, S.-Y. Meng, and L.-A. Wu, *Opt. Lett.* **42**, 5290 (2017).
- ¹⁰N. Radwell, K. J. Mitchell, G. M. Gibson, M. P. Edgar, R. Bowman, and M. J. Padgett, *Optica* **1**, 285 (2014).
- ¹¹Y. Wu, P. Ye, I. O. Mirza, G. R. Arce, and D. W. Prather, *Opt. Express* **18**, 24565 (2010).
- ¹²C. Zhao, W. Gong, M. Chen, E. Li, H. Wang, W. Xu, and S. Han, *Appl. Phys. Lett.* **101**, 141123 (2012).
- ¹³B. I. Erkmen, *J. Opt. Soc. Am. A* **29**, 782 (2012).
- ¹⁴W. Gong and S. Han, *Sci. Rep.* **5**, 9280 (2015).
- ¹⁵R. S. Bennink, S. J. Bentley, and R. W. Boyd, *Phys. Rev. Lett.* **89**, 113601 (2002).
- ¹⁶R. I. Khakimov, B. M. Henson, D. K. Shin, S. S. Hodgman, R. G. Dall, K. G. H. Baldwin, and A. G. Truscott, *Nature* **540**, 100 (2016).
- ¹⁷S. Li, F. Cropp, K. Kabra, T. J. Lane, G. Wetzstein, P. Musumeci, and D. Ratner, *Phys. Rev. Lett.* **121**, 114801 (2018).
- ¹⁸C. M. Watts, D. Shrekenhamer, J. Montoya, G. Lipworth, J. Hunt, T. Slesman, S. Krishna, D. R. Smith, and W. J. Padilla, *Nat. Photonics* **8**, 605 (2014).
- ¹⁹B. Sun, M. P. Edgar, R. Bowman, L. E. Vittert, S. Welsh, A. Bowman, and M. J. Padgett, *Science* **340**, 844 (2013).
- ²⁰M. J. Sun, M. P. Edgar, G. M. Gibson, B. Sun, N. Radwell, R. Lamb, and M. J. Padgett, *Nat. Commun.* **7**, 12010 (2016).
- ²¹A. M. Kingston, D. Pelliccia, A. Rack, M. P. Olbinado, Y. Cheng, G. R. Myers, and D. M. Paganin, *Optica* **5**, 1516 (2018).
- ²²J. H. Shapiro, *Phys. Rev. A* **78**, 061802 (2008).
- ²³Y. Bromberg, O. Katz, and Y. Silberberg, *Phys. Rev. A* **79**, 053840 (2009).
- ²⁴D. Pelliccia, M. P. Olbinado, A. Rack, A. M. Kingston, G. R. Myers, and D. M. Paganin, *IUCrj* **5**, 428 (2018).
- ²⁵A. Schori, D. Borodin, K. Tamasaku, and S. Shwartz, *Phys. Rev. A* **97**, 063804 (2018).
- ²⁶D. Pelliccia, A. Rack, M. Scheel, V. Cantelli, and D. M. Paganin, *Phys. Rev. Lett.* **117**, 113902 (2016).
- ²⁷A. Schori and S. Shwartz, *Opt. Express* **25**, 14822 (2017).
- ²⁸H. Yu, R. Lu, S. Han, H. Xie, G. Du, T. Xiao, and D. Zhu, *Phys. Rev. Lett.* **117**, 113901 (2016).
- ²⁹A.-X. Zhang, Y.-H. He, L.-A. Wu, L.-M. Chen, and B.-B. Wang, *Optica* **5**, 374 (2018).
- ³⁰Y. Klein, A. Schori, I. P. Dolbnya, K. Sawhney, and S. Shwartz, *Opt. Express* **27**, 3284 (2019).
- ³¹M. F. Li, L. Yan, R. Yang, and Y. X. Liu, *Acta Phys. Sin.* **68**, 064202 (2019) (in Chinese).
- ³²J. Romberg, M. F. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, K. F. Kelly, R. G. Baraniuk, T. Sun, K. F. Kelly, and R. G. Baraniuk, *IEEE Signal Process. Mag.* **25**, 83 (2008).
- ³³E. J. Candès, J. Romberg, and T. Tao, *IEEE Trans. Inf. Theory* **52**, 489 (2006).
- ³⁴O. Katz, Y. Bromberg, and Y. Silberberg, *Appl. Phys. Lett.* **95**, 131110 (2009).
- ³⁵C. Li, W. Yin, H. Jiang, and Y. Zhang, *Comput. Optim. Appl.* **56**, 507 (2013).
- ³⁶Y. Lecun, Y. Bengio, and G. Hinton, *Nature* **521**, 436 (2015).
- ³⁷M. Lyu, W. Wang, H. Wang, H. Wang, G. Li, N. Chen, and G. Situ, *Sci. Rep.* **7**, 17865 (2017).
- ³⁸Y. Zhang, Y. Rivenson, H. Günaydin, H. Wang, A. Ozcan, and Z. Göröcs, *Optica* **4**, 1437 (2017).
- ³⁹A. Goy, K. Arthur, S. Li, and G. Barbastathis, *Phys. Rev. Lett.* **121**, 243902 (2018).
- ⁴⁰P. Liu, H. Zhang, K. Zhang, L. Lin, and W. Zuo, [arXiv:1805.07071](https://arxiv.org/abs/1805.07071) (2018), also available as Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Work (IEEE, 2018), pp. 773–782.