UC Davis
UC Davis Previously Published Works

Title
Deep Compressed Imaging via Optimized-Pattern Scanning.

Permalink
https://escholarship.org/uc/item/137655b6

Journal
Photonics Research, 9(3)

ISSN
2327-9125

Authors
Zhang, Kangning
Hu, Junjie
Yang, Weijian

Publication Date
2021-03-01

DOI
10.1364/prj.410556

Peer reviewed

Deep Compressed Imaging via Optimized-Pattern Scanning

KANGNING ZHANG[1], **JUNJIE HU**[1], **WEIJIAN YANG**[1,*]

[1]Department of Electrical and Computer Engineering, University of California, Davis, CA 95616, USA

Abstract

The need for high-speed imaging in applications such as biomedicine, surveillance and consumer electronics has called for new developments of imaging systems. While the industrial effort continuously pushes the advance of silicon focal plane array image sensors, imaging through a single-pixel detector has gained significant interests thanks to the development of computational algorithms. Here, we present a new imaging modality, Deep Compressed Imaging via Optimized-Pattern Scanning (DeCIOPS), which can significantly increase the acquisition speed for a single-detector-based imaging system. We project and scan an illumination pattern across the object and collect the sampling signal with a single-pixel detector. We develop an innovative end-to-end optimized auto-encoder, using a deep neural network and compressed sensing algorithm, to optimize the illumination pattern, which allows us to reconstruct faithfully the image from a small number of samples, and with a high frame rate. Compared with the conventional switching-mask based single-pixel camera and point scanning imaging systems, our method achieves a much higher imaging speed, while retaining a similar imaging quality. We experimentally validated this imaging modality in the settings of both continuous-wave (CW) illumination and pulsed light illumination and showed high-quality image reconstructions with a high compressed sampling rate. This new compressed sensing modality could be widely applied in different imaging systems, enabling new applications which require high imaging speed.

1. INTRODUCTION

High-speed imaging has become more and more crucial in many new applications, such as in biomedicine, surveillance, and consumer electronics. There are two roadmaps for high-speed optical imaging: engineering a faster focal plane array image sensor and developing new imaging modalities using a single-pixel detector. Although intense industrial efforts have been made, high-speed and low-noise silicon focal plane array cameras are still expensive. Furthermore, imaging at wavelengths outside the silicon sensitivity spectrum can make the focal plane array cameras considerably more complicated [1]. In contrast, imaging through a single-pixel detector, which shrinks a photodetector array down to a single unit [2–9], can enormously reduce the cost and offer additional features such as reduced pixel crosstalk. A popular imaging modality of single-pixel detector is based on point scanning [5–9] (Fig. 1a), for example, the laser-scanning microscopes that are commonly used in

---

biomedicine [5–8]. However, such methods are speed limited due to the point-by-point data acquisition. Another approach with single-pixel detector relies on compressed sensing (CS) [10–12], represented by the switching-mask based single-pixel cameras [2–4] (Fig. 1b). There, the images are encoded by a series of spatially well-designed sampling patterns. For each pattern, all pixels across the entire image are summed and collected by the detector. Leveraging the general prior knowledge of sparsity in images, CS is used to reconstruct the image through a small number of measurements [2–4, 13, 14]. Although the sampling rate can be below the Nyquist criterion, the imaging frame rate is limited by how fast the sampling pattern can be switched and cycled, which is typically conducted by a digital micromirror device (DMD) [15] and operates <22.7 kHz. Using LED array to generate the pattern could increase the overall speed, but so far only 32×32 pixel images have been demonstrated [16] and it may be expensive to scale up. It does not allow passive light illumination (i.e. structured detection) either.

In this paper, we propose and demonstrate a deep compressed sensing modality, which can significantly increase the imaging speed while preserving a high reconstruction quality. This approach combines the strength of both compressed sensing and point-scanning imaging, and we term it Deep Compressed Imaging via Optimized-Pattern Scanning (DeCIOPS). Instead of projecting multiple binary patterns onto the entire object sequentially, we utilize only one gray-scale optimized pattern and project it to a small subset of the object. We then scan the pattern across the object by using fast scanning mirrors and collect the signal convolutionally using a single-pixel detector (Fig. 1c). Compared with the conventional single-pixel camera which relies on sequentially switching the sampling pattern on a DMD, our scanning approach significantly increases the sampling speed. Compared with the point-scanning system, our method samples a much larger portion of the object at once and recover the resolution computationally. This allows a great reduction of the sampling number and thus increases the frame rate. We note that the improvement of imaging speed does not require an increase of light energy. In fact, the required light dosage in our method is smaller than the conventional point-scanning system due to a reduced number of measurements. We build an auto-encoder framework [17] to optimize the sampling pattern. The image acquisition system is treated as an encoder, where the high-resolution object is encoded through the sampling pattern into a few measurements. We then formulate an ISTA-Net [18], a CS-induced neural network inspired by the Iterative Shrinkage-Thresholding Algorithm (ISTA) [19], as a decoder to reconstruct the image. This auto-encoder is trained in an end-to-end fashion. Such a framework can learn an optimized sampling pattern and simultaneously recover a high-resolution image by extracting the feature of sparsity and searching the optimal pair of encoder and decoder with the lowest incoherence [11], which is one key feature of DeCIOPS versus other compressed sensing or deep learning-based super-resolution imaging modalities [20–23]. This new imaging modality can be accustomed to any light scanning imaging system and will greatly benefit the high-throughput imaging applications.

This paper is organized as below. In Section 2, we introduce the mechanisms of the imaging modality in two configurations of illuminations – continuous-wave (CW) and pulsed light source, as well as the auto-encoder framework and the deep compressed sensing neural network for optimizing the imaging and reconstruction. In Section 3 and Section 4, we show

the simulation results and experimental results. In Section 5, we discuss the system performance under different signal-to-noise ratio and compression ratio, and how DeCIOPS can be applied in two-photon microscopy and passive lighting condition.

2. PRINCIPLE

A. Image formation

In DeCIOPS, an illumination pattern is generated through a mask, and is scanned across the object by using a set of scanning mirrors. The detector records a subsample of the 2D convolution between the pattern and the object (Fig. 2). Similar to any point-scanning imaging system, DeCIOPS can use either CW or pulsed light sources. The former is commonly used in imaging systems, whereas the latter is specialized for nonlinear microscopy. When a CW light source or high-repetition-rate pulsed light source is used, the detector continuously integrates the signal as the pattern scans. Hence, in DeCIOPS, we project a rectangle-shape pattern to the object. By finely adjusting the integration time, each acquisition measures a square subset of the object with a desired resolution (Fig. 2c). In case of a low-repetition-rate pulsed light source, the sampling time stamps of the detector is synchronized with the pulse train, and a square-shape pattern is projected to the object. By matching the spatial sampling step with the size of the pattern, the entire object is sampled appropriately (Fig. 2d).

Mathematically, in low-repetition-rate pulsed light illumination, the image formation of DeCIOPS can be expressed as

$$\boldsymbol{b} = f_{n \times n}(\boldsymbol{x} * \boldsymbol{g}), \tag{1}$$

where $\boldsymbol{x}$ is the object, $\boldsymbol{g}$ is a square shape illumination pattern, * represents the 2D convolution, $f_{n \times n}(\cdot)$ models the $n \times n$ undersampling, and $\boldsymbol{b}$ is the measured image. Here, we assume that the mask has a size of $n \times n$ pixels. In the particular case where $\boldsymbol{g}$ is a uniform mask g_1 (Fig. 2d, left), Eq. (1) is equivalent to a naïve undersampling by unweighted averaging every $n \times n$ pixels of the full resolution image (Fig. 2a) acquired in single-point scanning. As discussed in Section 2.B, $\boldsymbol{g}$ can be optimized to achieve the best image reconstruction performance (Fig. 2d, right).

In the CW light or high-repetition-rate pulsed light illumination case, where the detector continuously integrates the signal, we configure the illumination pattern in a size of $n \times 1$ pixels. When this pattern sweeps n columns, the information of $n \times n$ pixels is integrated into a single measurement (Fig. 2c). We can use the same mathematical formulation as Eq. (1) to model the image formation, where each column in the $n \times n$ mask $\boldsymbol{g}$ is identical.

B. End-to-end optimized auto-encoder framework

In DeCIOPS, we build an auto-encoder framework to simultaneously learn the optimized mask pattern and a neural network for image reconstruction (Fig. 3). The encoder block models the image formation through the following expression:

$$b^+ = f_{n \times n}(x * g) + \epsilon = F(x) + \epsilon = \Phi x + \epsilon, \tag{2}$$

where F is an operator and Φ is the linear transfer matrix, both equivalent to the subsampled convolution with the mask g in Eq. (1), ϵ is the additive noise inherent in the imaging system, and b^+ is the measured mask-encoded image.

The decoder takes b^+ as the input and aims to reconstruct the original object x, by solving the following convex optimization problem with a constraint of the sparse representation of x:

$$\underset{x}{\operatorname{argmin}} \frac{1}{2} \left\| \Phi x - b^+ \right\|_2^2 + \lambda \|\Psi(x)\|_1, \tag{3}$$

where $\Psi(x)$ denotes a transform of x into a sparse representation under the basis of Ψ, and λ is a hyperparameter.

The solution of the problem in Eq. (3) can be initialized by calculating the pseudo-inverse of the encoder from the measurement b^+. We then use the ISTA algorithm [19] to find an optimized solution of x by iterating the following two steps:

$$r^{(k)} = x^{(k-1)} - \rho \Phi^T \left(\Phi x^{(k-1)} - b^+ \right), \tag{4}$$

$$x^{(k)} = \underset{x}{\operatorname{argmin}} \frac{1}{2} \left\| x - r^{(k)} \right\|_2^2 + \lambda \|\Psi(x)\|_1, \tag{5}$$

where k denotes the kth ISTA iteration step and ρ is the step size.

As Ψ is predefined empirically and may not be suitable for the data, we adopt ISTA-Net algorithm [18] which can learn Ψ through the data. In ISTA-Net, Ψ is replaced with a trainable neural network $\mathcal{F}(\cdot)$, and the optimization problem in Eq. (3) can be rephrased into the following L1-norm regularization problem with a nonlinear transform $\mathcal{F}$

$$\underset{x}{\operatorname{argmin}} \frac{1}{2} \left\| \Phi x - b^+ \right\|_2^2 + \lambda \|\mathcal{F}(x)\|_1. \tag{6}$$

The kth iteration step in the original ISTA algorithm is replaced by a series of symmetric learnable parameters in the kth ISTA-Net phase:

$$\mathcal{F}\left(x^{(k)}\right) = soft\left(\mathcal{F}\left(r^{(k)}\right), \theta\right), \tag{7}$$

$$x^{(k)} = \mathcal{F}^{-1}\left(soft\left(\mathcal{F}\left(r^{(k)}\right), \theta\right)\right), \tag{8}$$

where θ is a learnable parameter in the kth module, $\mathscr{F}^{-1}(\cdot)$ is the inverse of $\mathscr{F}(\cdot)$, and $Soft(\cdot)$ represents soft shrinkage threshold. Finally, we obtain the output of the decoder x^N after a total number of N ISTA-Net phases.

The loss function of ISTA-Net is obtained by calculating the mean-square-error (MSE) between the output and the ground truth with the constraint of $\mathscr{F}^{-1} \circ \mathscr{F} = I$ as both $\mathscr{F}^{-1}$ and $\mathscr{F}$ are learnable and symmetric, where I is the identity operator. As a result, we have the following loss function with the symmetry constraint:

$$L_{total} = L_{error} + \gamma L_{constraint}$$

$$= \left\| x^{(N)} - x \right\|_2^2 + \gamma \left(\sum_{k=1}^{N} \left\| \mathscr{F}^{-1(k)}\left(\mathscr{F}^{(k)}(x)\right) - x \right\|_2^2 \right), \tag{9}$$

where γ is the weight of the symmetry constraint.

3. SIMULATION RESULTS

We trained the auto-encoder using 1500 samples of natural scenes (2D grey-scale image, 256×256 pixel size) from ImageNet [24], and validated the model with 79 samples from two widely used benchmark datasets: Set11 [25] and BSD68 [26]. As an illustration, we chose an undersampling rate of 6.25% (4×4 undersampling), and initialized the pattern g as a 4×4 normalized random Gaussian matrix. γ was set to be 0.01 in the loss function, accompanied by Adam optimization with a learning rate of 1×10^{-4}. We included additive noises in the measurement (5~10% of the signal strength). The training was performed on a GPU RTX2080Ti 11GB. The training work of N=9 ISTA-Net phases takes ~5 hours for 200 epochs with a batch size of 5. We trained two independent auto-encoders, one with a constraint on g so each column of g is identical (CW light or high-repetition-rate pulsed light illumination), and one without such a constraint (low-repetition-rate pulsed light illumination). The reconstruction results are evaluated in terms of Peak Signal-to-Noise Ratio (PSNR) and spatial resolution by using Fourier ring correlation [27] in the validation dataset. As a control measure, we compared the reconstruction performance of the optimized pattern (Fig. 2c–d, right) with a random pattern, and two naïve undersampling schemes of the full resolution image either through an unweighted averaging of 4×4 pixels (equivalent to the uniform pattern, Fig. 2c–d, left) or a simple dropout (i.e. pick one pixel in every 4×4 and drop out the others, Fig. 2b). All the simulation groups employ an independently trained ISTA-Net for image reconstruction. In addition, we also used a B-spline interpolation [28] to reconstruct the image undersampled through simple dropout.

We compared the measurements and their corresponding reconstruction results for various schemes (Fig. 4). In all cases, the reconstructed results can resolve higher resolution features than the raw measurement. In the undersampling through simple dropout (Fig. 4a), ISTA-Net shows better performance than the interpolation. These results demonstrate the effectiveness of ISTA-Net. As the undersampling through dropout (Fig. 4a) misses a

substantial amount of information in the original object, its reconstruction result is expected to be the worst. Comparing the undersampling through the uniform pattern (Fig. 4b) and the optimized pattern (Fig. 4c–d), we find that the optimized pattern cases show reconstruction results with sharper edges. We reason this as the optimized pattern balances both the high-frequency and low-frequency components of the original object during sampling; whereas the uniform pattern performs a lowpass filtering such that the high-frequency component is lost before reconstructed by the decoder. Indeed, when comparing the quantitative results of the PSNR (Fig. 4e) and the resolution of the reconstructed image through Fourier ring correlation (Fig. 4f), we find the optimized pattern case shows the best performance. It is important to note that the optimized pattern also outperforms the exemplary random pattern (Fig. 4c–d) for both PSNR and resolution. This verifies the effectiveness of our end-to-end optimized auto-encoder framework. We also note that the optimized mask without constraints of identical column show a better performance in resolution than the one with constrains, though their PSNR does not show significant difference.

4. EXPERIMENTAL RESULTS

To validate the numerical simulation results, we built an imaging system for DeCIOPS (Fig. 5, more details in Appendix 1). Here, we used a DMD to generate the light pattern as it offers a great flexibility in comparing the performance between different patterns. In general, as DeCIOPS requires only one illumination pattern, a fixed pattern mask can be used. To generate a gray-scale mask from the DMD binary pixels, we binned 32×32 pixels in the DMD into a super-pixel which could provide up to 1025 gray-scale levels. A total of 4 $\times$ 4 super-pixels were programmed to generate the optimized mask pattern. A 520 nm diode CW laser source was expanded in beam size and collimated to illuminate the mask pattern on the DMD. The spatially encoded light then passed a 4f system composed of a tube lens and objective lens to reduce its beam size. The light pattern was then raster scanned across the sample through an optically coupled resonant-galvo scanner set. We used a photolithography mask as the sample. The light transmitting through the sample was then collected by a photodetector through a collection lens. While we built this transmission-based imaging system for simplicity, we could turn it into a reflection-based system through a beam splitter in front of the sample. We used ScanImage [29] as control software for data acquisition. By adding another 4f system composed of cylindrical lenses right after the objective lens, we could turn the 4×4 size pattern into 4×1 size (Appendix 1.A). Thus, we could use the same setup to validate the different sampling schemes.

As the optical mode from the diode laser did not have uniform intensity and the pattern could be corrupted by laser interference, we calibrated the DMD to ensure the illumination pattern on the sample plane matched well with the design (Appendix 1.C). While our imaging system is naturally a CW light imaging system, we could also mimic the pulsed light source condition through an additional digital sampling step after the image acquisition (Appendix 1.E).

A. Reconstruction results with a CW light source

We evaluated the experimental results in the naturally CW light source setting with the 4×1 pattern mask. We compared the reconstruction results across the four undersampling schemes: simple dropout, the uniform illumination pattern, an exemplary random illumination pattern, and the optimized illumination pattern for various samples (Fig. 6). The ground truth images were obtained by the high-resolution point-scanning approach. Though at an undersampling rate of 6.25%, all reconstructed results through ISTA-Net show a significant improvement from the raw measurement where high-resolution features are better resolved. Comparing between different illumination patterns, we notice more details on the edges of the reconstructed images in the optimized-pattern illumination cases (Fig. 6a–d). Using the metric of PSNR (Fig. 6e) and spatial resolution (Fig. 6f), we find that DeCIOPS outperforms the simple dropout, the uniform illumination pattern and the random illumination pattern with an average of 2.41 dB, 1.82 dB and 1.73 dB improvement in PSNR, and an average of 18.9%, 14.0% and 11.6% improvement on spatial resolution respectively.

B. Reconstruction results with a low-repetition-rate pulsed light source

Using the same experimental setup and CW light source, we preprocess the acquired data through a digital undersampling process (Appendix 1.E) so the results mimic that acquired by using a low-repetition-rate pulsed light (Fig. 7). As the integration time of each measured pixel is reduced, the raw measurements have a lower signal-to-noise ratio (SNR) compared with the CW light source setting. Nevertheless, the reconstruction results show a greatly improved quality. On average, the PSNR of the reconstructed images using the optimized illumination pattern shows an improvement of 1.73 dB, 1.41 dB and 0.64 dB when they are compared with the simple dropout, uniform illumination pattern and the exemplary random illumination pattern, respectively. The improvement in the spatial resolution in the optimized illumination pattern against the simple dropout, the uniform and the random illumination pattern is 14.06%, 10.33% and 7.95%, respectively. Compared with the CW light source setting, the performance improvement of the optimized illumination pattern is reduced because of a reduced SNR in the raw measurement in the low-repetition-rate pulsed light source setting. In the Session 5.A, we further discuss how the SNR influences the reconstructed results.

5. DISCUSSION

A. The influence of SNR on image reconstruction

In this section, we study the influence of SNR of the raw measurement on the PSNR and pixel resolution of the reconstruction (Fig. 8). We performed simulations by adding different noise levels in the measurement. For each SNR level, we trained the DeCIOPS independently. As expected, both the PSNR and pixel resolution of the reconstruction improve as the SNR increases and saturates at a high SNR. Compared with the naïve undersampling through simple dropout (blue curve, Fig. 8) and unweighted average (green curve, Fig. 8), the performance of the optimized illumination pattern (red curve, Fig. 8) has a larger improvement in the PSNR and pixel resolution as SNR increases. This phenomenon emphasizes the advantage of the DeCIOPS, where the optimal encoder and the decoder are

able to match each other better than other independent untrained encoders. As the noise reduces in the raw measurement (i.e. increasing SNR), the network tends to pay more attention on the reconstruction rather than denoising. As a result, the performance difference among different illumination patterns increase.

B. Compressed ratio and size of the optimized pattern

In the above simulations and experiments, we set the illumination pattern size such that each pixel in the original high-resolution object is measured only once during the scanning. We term this as the matching condition between the pattern and undersampling rate. Here, for each specific undersampling rate, we simulated the PSNR of the reconstructed images versus the size of the illumination pattern (Fig. 9). We notice that DeCIOPS generates the best reconstruction results when the size of the pattern matches to the two-dimensional undersampling rate, where the imaging modality in the encoder is a linear orthonormal transform algebraically. When the size of the pattern is smaller than that of the matching condition, there appears a sharp drop in the reconstruction performance. This is attributed to the permanent information loss from the unsampled pixels in between the adjacent measurements, such that the sensing basis Φ^T in Eq. (2) forms a singular transform matrix and thus degrades the quality of reconstruction. When the size of the pattern is larger than that of the matching condition, the sensing basis Φ^T becomes less column orthonormal, resulting in a less accurate reconstruction under the CS framework and thus a drop of the PSNR in the reconstructed images.

C. Optimized pattern versus random pattern

In the conventional switching-mask based single pixel camera, random mask is one of the commonly used sampling basis as it is incoherent with the spatial property of the sample. We found the same in the pattern scanning scheme: the random pattern shows a superior performance compared with the uniformed pattern. However, the optimized pattern, found by the auto-encoder through the end-to-end optimization, outperforms the random pattern. This is expected as random pattern attained by Monte Carlo method is independent to the dataset of specific task unlike the optimized pattern. Furthermore, the performance improvement of the optimized pattern over the random pattern increases when the measurements are more highly undersampled, as seen in the comparison between 6.25% (4 × 4 pattern) and 1.5625% (8 × 8 pattern) (Table 1). This is attributed by more trainable parameters in case of the optimized pattern and more uncertain random variables in case of the random pattern.

D. Comparison with conventional switching-mask based single-pixel camera

As both DeCIOPS and the conventional switching-mask based single-pixel camera (Fig. 1b) leverages compressed sensing, we expect their reconstruction quality is similar at the same undersampling rate. Indeed, the one-way ANOVA test shows no significant difference for PSNR and pixel resolution in the reconstructed images between these two approaches, for both simulation and experiment (Fig. 10, Table 2, 256 × 256 pixels high-resolution objects, 6.25% undersampling rate, ISTA-Net reconstruction framework for both DeCIOPS and switching-mask based single-pixel camera). However, DeCIOPS has a much faster acquisition speed. To image a high-resolution object with 256 × 256 pixels, DeCIOPS

acquires 64×64 measurements at an undersampling rate of 6.25%. By using a scanning system with an 8 kHz resonant scanner, the acquisition time is 4 ms. In the conventional switching-mask based single-pixel camera, at the same undersampling rate, it takes 180 ms to cycle 64×64 = patterns using a 22.7 kHz high-speed DMD, without considering the integration time of the detector. Thus, the image speed of our method is orders of magnitude faster than the conventional compressed sensing approach using a single-pixel camera.

E. Advantage of ISTA-Net in image reconstruction

In DeCIOPS, we apply ISTA-Net as the decoder. ISTA-Net is a CS-induced neural network. Compared with the conventional optimization algorithms where the regularization term is designed empirically, ISTA-Net is entirely data-driven and can learn an optimized regularization through the neural network. Compared with other neural networks that could be used for super-resolution, such as U-Net [30, 31] and Densely Connected Super-Resolution Network (DCSRN) [32, 33], the embedded CS algorithm in ISTA-Net fits better in the motivation of DeCIOPS and other compressed sensing framework, i.e. performing fewer measurements while being able to reconstruct high resolution image. Indeed, when we compare the PSNR and pixel resolution of the reconstructed images across ISTA-Net, U-Net and DCSRN, ISTA-Net shows the best performance (Appendix 2). In addition to DeCIOPS, we believe ISTA-Net could also benefit other applications such as denoise [34], fast Magnetic Resonance Imaging (MRI) [35], and other super-resolution imaging modalities [18].

F. Advantage of end-to-end optimized auto-encoder and its application in future imaging systems

In most existing optical imaging modalities, the image formation is empirically designed and optimized, and the deconvolution or object reconstruction algorithm is subsequently tailored to the image formation process. The recent development of low-cost, advanced micro-optics manufacturing techniques such as 3D printing and micro-nano-fabrication [36–40] allows rapid prototyping of user-designed optical elements, which opens new opportunities to redesign the image formation process that best fits the specific applications. Instead of sequentially designing the image formation and the reconstruction algorithm, their joint end-to-end optimization produces a global optimal solution [41–43], which is the underlying principle of DeCIOPS. We use an auto-encoder to model the image formation and reconstruction within a single framework and perform an end-to-end training to optimize the sampling pattern and ISTA-Net simultaneously. Our results show the optimized sampling pattern indeed results in the best overall performance. Such an end-to-end training and data-driven approach prevents any empirical bias that may negatively impact the design. We envision that such an approach will enable many challenging applications such as super-resolution imaging [23, 42–46], 3D imaging [41, 47–49], and high-speed computational camera [50–52].

G. Applicability in two-photon microscopy

While our imaging system used a CW light source, we mimicked the experimental condition of a pulsed light source and successfully demonstrated the applicability and excellent performance of DeCIOPS. This opens a new avenue to apply DeCIOPS in two-photon

microscopy. In conventional two-photon microscopes [5, 7, 8], the image is acquired through pixel-by-pixel point scanning. While this enables deep tissue imaging as it resists light scattering, it reduces the imaging speed. Recently, there have been multiple reports applying compressed sensing in two-photon microscopy, with the same approach in the conventional switching-mask based single-pixel camera [53–55]. However, the improvement on the imaging speed is limited due to the low switching speed of DMD. When applying DeCIOPS in two-photon microscopy, we expect that our approach will significantly increase the imaging speed and will notably benefit functional imaging through two-photon microscopy.

H. Passive light illumination

In our experiment, DeCIOPS is implemented using active light illumination (i.e. structured illumination), which is commonly used in biomedical imaging. In other imaging systems, passive light illumination may be preferred. In fact, any passive light illumination wide-field imaging using a focal plane array (i.e. camera) can be converted to DeCIOPS (Appendix 3). A scanner can be added to the passive wide-field imaging to scan the entire image originally projected to the camera. By inserting a fixed mask with an appropriate aperture at the plane where the image is scanned, a single-pixel detector can measure the subsampled convolution between the mask and the original image. The object can then be reconstructed using the same algorithms in DeCIOPS for the active light illumination cases. This way, a structured detection version of DeCIOPS can be implemented.

6. CONCLUSION

We demonstrated a new high-speed imaging modality, DeCIOPS, by synthesizing the strength of conventional point scanning and single-pixel camera through compressed sensing. The high-speed imaging arises from the fast beam scanning mechanism and a highly efficient sampling scheme through compressed sensing; meanwhile, an auto-encoder framework allows the simultaneous optimization of the image formation and reconstruction process in DeCIOPS. We validated DeCIOPS through both simulation and experiments, in both CW and pulsed light source conditions. This new image modality can be adapted to any existing imaging systems using beam scanning, such as confocal microscope and two-photon microscope, or wide-field cameras with an added scanning system and will benefit broad applications requiring high speed imaging.

Acknowledgments.

APPENDIX 1.: Experimental setup of DeCIOPS

A. Optical setup

The optical setup of DeCIOPS illustrated in Fig. 5 scans an $n \times n$ pattern on the object and is suitable for low-repetition-rate pulsed light illumination setting. For CW light or high-repetition-rate pulsed light illumination, an $n \times 1$ pattern is scanned. When this pattern

sweeps n columns, the information of $n \times n$ pixels can be integrated into a single measurement. To generate the $n \times 1$ pattern, we set the column to be identical in the $n \times n$ pattern and add a 4f system composed of cylindrical lenses after the objective lens. The $n \times n$ pattern is then shrunk in one dimension by a factor of n into $n \times 1$ pattern (Fig. 11). The parameters of the lenses used in the setup are listed in Table 3.

B. Pattern generation

The pattern is generated from a DMD illuminated by a parallel beam. In our experiment, the pattern has 4×4 pixels, each of which has a gray-scale throughput/intensity ranging from 0 to 1. As each DMD mirror pixel only provides a binary (on or off) light throughput, we group 32×32 DMD mirror pixels as one super pixel, so a gray-scale light throughput becomes feasible within one super-pixel. Each super-pixel thus provides 1025 states in the range from 0 to 1. The multistate light throughput is achieved by randomly selecting parts of the mirrors to be on, the number of which is the product of the desired light throughput and total number of mirror pixels (1024).

In the case of 4×1 pattern scanning, another 4f system composed of cylindrical lenses is added after the objective lens (Fig. 11). Such a system shrinks the 4×4 pattern in one dimension by a factor of four.
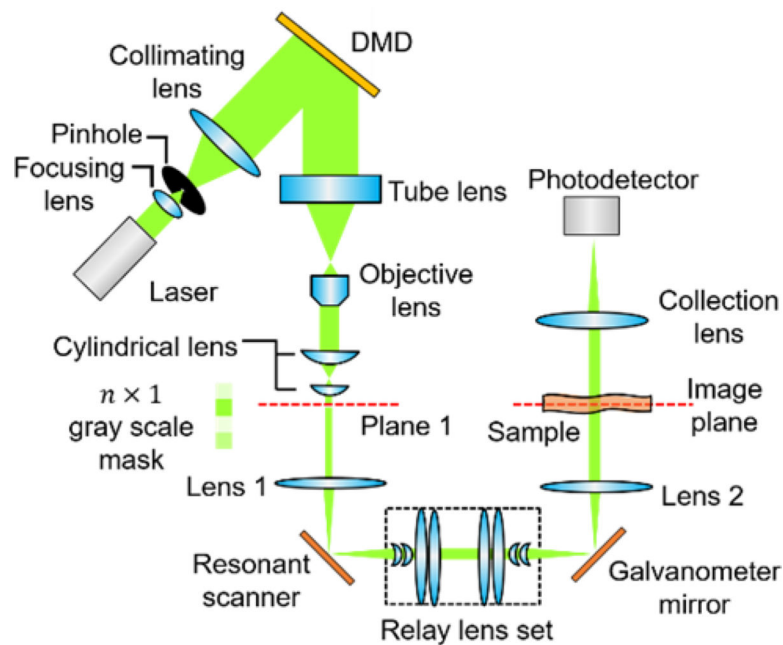

Fig. 11.
The experimental setup of DeCIOPS that generates an $n \times 1$ size pattern and scans it across the sample. The setup is similar as that generates the $n \times n$ size pattern shown in Fig. 5, but with a 4f system composed of cylindrical lenses added after the objective lens to shrink the original $n \times n$ size pattern in one dimension by a factor of n into $n \times 1$ size. The red dashed line (plane 1 and object plane) indicates the conjugate plane of the gray-scale pattern mask.

To obtain the high-resolution ground truth image, we reconfigure the imaging system for point scanning. We replace the collimating lens between the pinhole and the DMD to the one with shorter focal length and set the displayed pattern on the DMD to be uniform. This reduces the beam size on the image plane, and the pattern effectively turns into a spot matching the size of a single super-pixel of the 4×4 pattern. This ensures the same overall energy collected by the photodetector and thus a similar SNR per measurement as in DeCIOPS.

C. Characterization of the pattern

As the optical mode from the diode laser was not symmetric, we implemented a spatial filter using a pinhole after the diode laser output. Nevertheless, there still appeared non-uniform intensity of the light illuminated on the DMD. Furthermore, the laser interference could corrupt the pattern. We thus fine-tuned the DMD pixel value to calibrate the intensity of the super pixels, so their values were as close to the design as possible.

We used an iterative approach to calibrate the intensity of the super pixel. The design pattern M_0 with the gray scale super-pixel, which was also written as $M^{(0)}$ for consistency, was first converted to a pattern $W^{(0)}$ on the DMD through an operator $\mathcal{D}$. $\mathcal{D}$ essentially converted each super-pixel in M_0 into 32×32 pixels in binary values on the DMD, with the on and off pixels randomly distributed. $\mathcal{D}^{-1}$ is the inverse operation. We used a camera to capture the projected pattern $M^{(1)}$ on the object plane. We then compared $M^{(1)}$ and M_0, and updated $W^{(0)}$ into $W^{(1)}$ through the following algorithm:

$$W^{(1)} = \mathcal{D}\left[\left(M^{(1)} - M_0\right) \times a + M^{(0)}\right],\qquad(10)$$

where a is the step size to control the update rate. For the Kth iteration, we have

$$W^{(k)} = \mathcal{D}\left[\left(M^{(k)} - M_0\right) \times a + \mathcal{D}^{-1}\left(W^{(k-1)}\right)\right],\qquad(11)$$

In general, the pattern could be calibrated well with <5 iterations. With this calibration, we verified that the projected pattern on the sample matched the designed pattern in most of the cases (Fig. 12). In the special case of the 4×1 random pattern, we took an alternative approach. We randomly generated the pattern and measured it at the object plane. We then used this pattern as the "design" to train the decoder for object reconstruction.
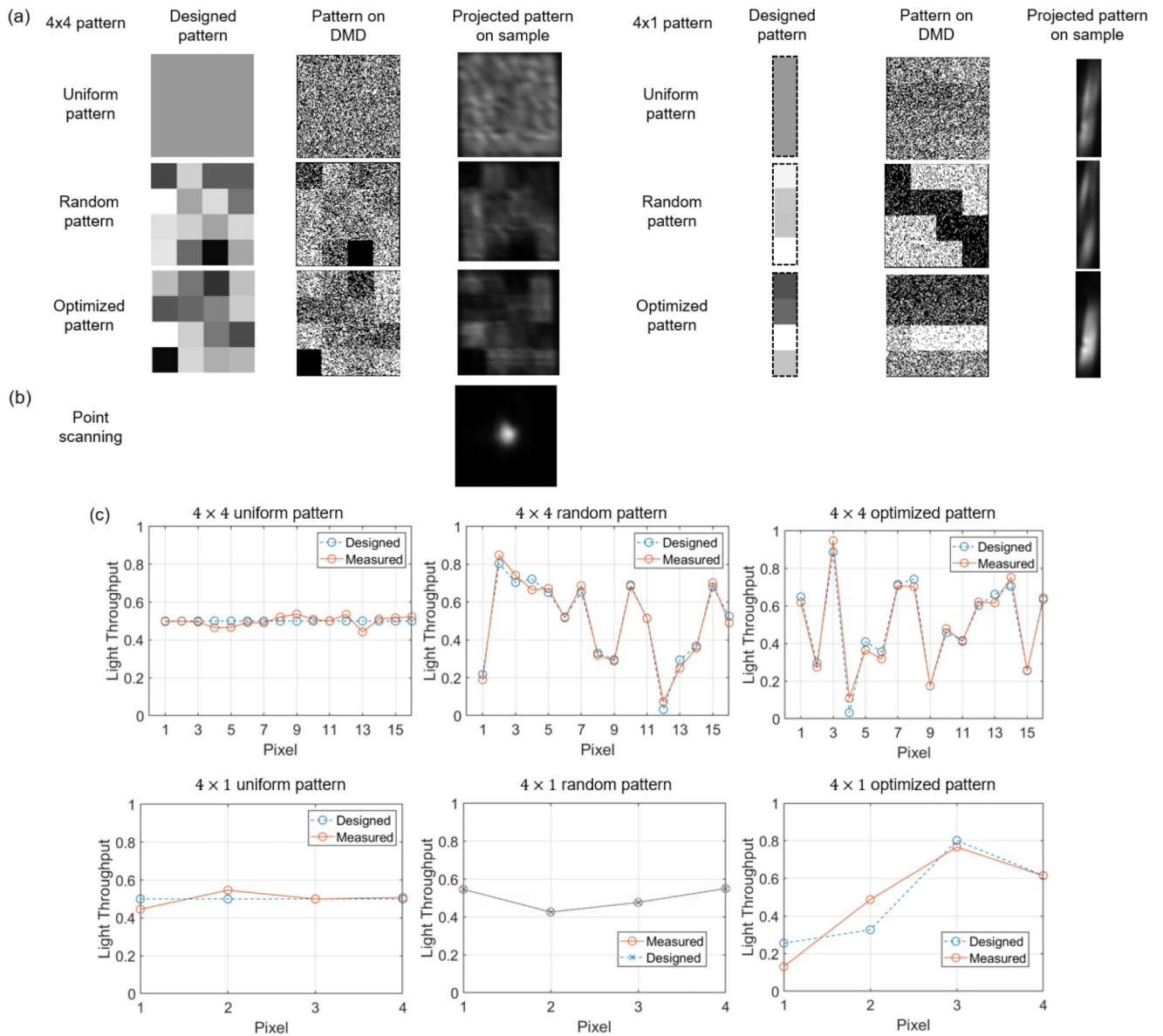
Fig. 12.
(a) The measured patterns on the sample (super-pixels) match well with the designed patterns. Each gray-scale super-pixel is generated by 32×32 binary pixels in the DMD. The left panel shows the cases for 4×4 patterns, and the right panel shows the cases for 4×1 pattern. (b) A single spot pattern is generated for conventional point scanning imaging to obtain the high-resolution ground truth of the sample. The spot size matches the size of a super-pixel. (c) Pixel-by-pixel comparisons between the measured patterns on the sample and the designed patterns show excellent matchings between the two.

D. Pixel resolution and field of view

The size of an individual DMD mirror is $13.6 \times 13.6 \ \mu m^2$, and thus the size of a super-pixel is $(32 \times 13.6) \times (32 \times 13.6) = 435.2 \times 435.2 \ \mu m^2$. The imaging system has a magnification of

4.44, resulting in a size of $\left(\frac{435.2}{4.44}\right) \times \left(\frac{435.2}{4.44}\right) = 97.9 \times 97.9 \mu m^2$ of a super-pixel on the image plane. A full resolution object with 256×256 pixels ($\sim 24.5 \times 24.5$ mm^2), requires a scanning range of $\pm 2.8^o \times \pm 2.8^o$ in the scanner set. We measured the illumination pattern on the object plane across different scanning angles (Fig. 13). The excellent field uniformity ensures the quality of DeCIOPS.
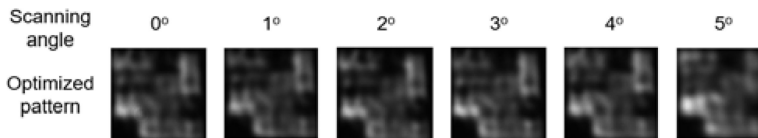

Fig. 13.
The measured patterns at the image plane stay consistent cross different scanning angles.

E. Data Acquisition

The data acquisition is performed using a high speed data acquisition card vDAQ and ScanImage software (Vidrio Technologies). As the illumination pattern continuously scans across each row, the data acquisition card samples the data from the photodetector at a rate higher than the single pixel rate. The data acquired within the duration of a single pixel is then automatically averaged/integrated and saved as a single pixel value. Compared with the full resolution single point scanning condition (256×256 pixels), we reduce the single pixel rate and the number of scanning lines by 75% in DeCIOPS (CW light setting), reaching an undersampling rate of 6.25% (64×64 pixels of measurement).

To mimic the low-repetition-rate pulsed light source condition, we sample the object with a high resolution at 2048×2048 pixels (corresponding to a high single pixel rate) and then downsample the acquired image digitally into 64×64 pixels by dropping all the other pixels. Here, each pixel has a small average/integration duration, and could thus be considered as being acquired by a single light pulse.

F. Estimation of signal-to-noise ratio

To calculate the signal-to-noise ratio of the image in the experiment, we acquire the same image for 20 times. For each pixel, we calculate the signal μ and the noise σ as the mean and the standard deviation across 20 measurements respectively. The signal-to-noise ratio for the pixel is then estimated as μ/σ. The signal-to-noise ratio of the entire image is taken as the average of the signal-to-noise ratio of all pixels.

APPENDIX 2.: Comparison between ISTA-Net, U-Net and DCSRN

We compare the performance of ISTA-Net, U-Net and DCSRN in object reconstruction. In the auto-encoder framework, the decoder implemented by ISTA-Net is replaced by U-Net or DCSRN. U-Net is widely used in image reconstruction and segmentation. It first condenses the size of the input images to extract its context and feature, and then grows them in an expanding path to perform local reconstruction [30]. DCSRN, derived from Densely Connected Convolutional Networks [33], has a faster training speed and accurate

reconstruction results, and is commonly used in applications such as 2D or 3D biomedical super-resolution imaging. In addition to ISTA-Net, U-Net and DCSRN, we used B-spline interpolation [28] to reconstruct the object undersampled through a simple dropout as a baseline. All the simulation was completed on GPU RTX1080Ti 11GB with 200 epochs and a batch size of 5. In each decoder except for B-Spline, we learned an optimized illumination pattern. We used the validation data set to evaluate the PSNR and pixel resolution of the reconstructed objects. U-Net, DCSRN and ISTA-Net all outperform B-spline interpolation. While U-Net and DCSRN do not show a significant difference in performance, ISTA-Net outperforms both U-Net and DCSRN with a ~1.85 dB increase in PSNR and 6.55 % improvement in resolution, at an undersampling rate of 6.25% (Fig. 14). The simulation results demonstrated a clear advantage of ISTA-Net, which is a Compressed Sensing (CS) inspired neural network, in DeCIOPS.
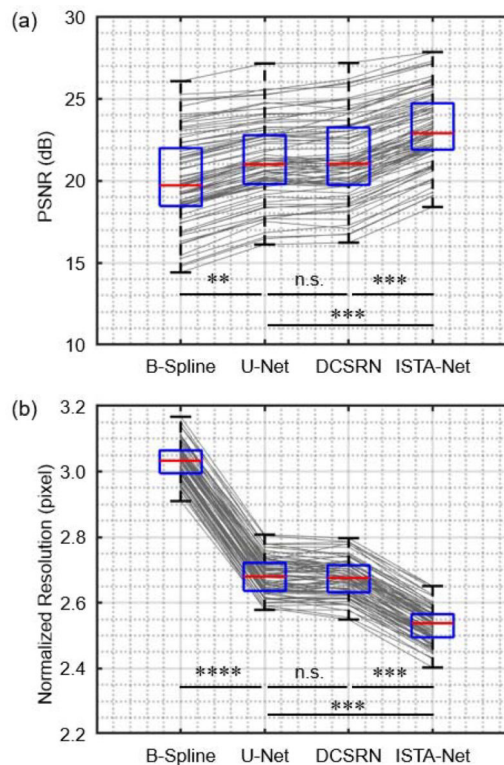

Fig. 14.
Comparison of (a) PSNR and (b) pixel resolution of the reconstructed objects of all 79 samples in the validation dataset for B-Spine, U-Net, DCSRN and ISTA-Net in the auto-encoder framework, at an undersampling rate of 6.25%. n.s., not significant; *, $p<0.05$; ***, $p<0.001$; ****, $p<0.0001$, in one-way Analysis of Variance (ANOVA).

APPENDIX 3.: DeCIOPS using passive light illumination

To implement DeCIOPS in a passive light illumination setting (i.e. structured detection), which is commonly used in photography, a scanner can be added to the passive wide-field imaging system to scan the entire image originally projected to the camera. By inserting a

fixed mask with an appropriate aperture at the plane where the image is scanned, a single pixel detector can measure the subsampled convolution between the mask and the original image (Fig. 15). The object can then be reconstructed using the same algorithms in DeCIOPS for the active light illumination cases.
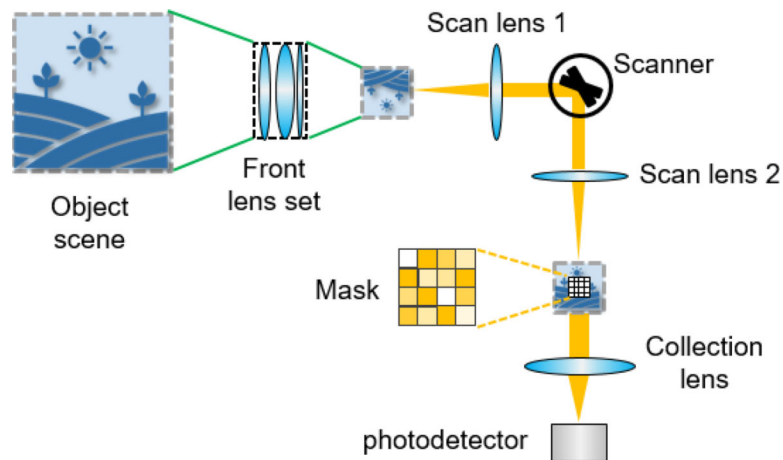

Fig. 15.
The optical setup of DeCIOPS with passive light illumination (i.e. structured detection), for applications such as photography.

References

1. Rogalski A, Infrared detectors (CRC press, 2010).

2. Duarte MF, Davenport MA, Takhar D, Laska JN, Sun T, Kelly KF, and Baraniuk RG, "Single-pixel imaging via compressive sampling," IEEE signal processing magazine 25, 83–91 (2008).

3. Edgar MP, Gibson GM, and Padgett MJ, "Principles and prospects for single-pixel imaging," Nature photonics 13, 13–20 (2019).

4. Gibson GM, Johnson SD, and Padgett MJ, "Single-pixel imaging 12 years on: a review," Opt Express 28, 28190–28208 (2020). [PubMed: 32988095]

5. Denk W, Strickler JH, and Webb WW, "Two-photon laser scanning fluorescence microscopy," Science 248, 73–76 (1990). [PubMed: 2321027]

6. Pawley JB, Handbook of biological confocal microscopy (Springer, 2006).

7. Zipfel WR, Williams RM, and Webb WW, "Nonlinear magic: multiphoton microscopy in the biosciences," Nat Biotechnol 21, 1369–1377 (2003). [PubMed: 14595365]

8. Helmchen F, and Denk W, "Deep tissue two-photon microscopy," Nat Methods 2, 932–940 (2005). [PubMed: 16299478]

9. Beiser L, "Fundamental architecture of optical scanning systems," Applied optics 34, 7307–7317 (1995). [PubMed: 21060601]

10. Candès EJ, "Compressive sampling," in Proceedings of the international congress of mathematicians(Madrid, Spain2006), pp. 1433–1452.

11. Candes E, and Romberg J, "Sparsity and incoherence in compressive sampling," Inverse problems 23, 969 (2007).

12. Candès EJ, and Wakin MB, "An introduction to compressive sampling," IEEE signal processing magazine 25, 21–30 (2008).

13. Spencer AP, Spokoyny B, Ray S, Sarvari F, and Harel E, "Mapping multidimensional electronic structure and ultrafast dynamics with single-element detection and compressive sensing," Nature communications 7, 1–6 (2016).

14. Bian L, Suo J, Dai Q, and Chen F, "Experimental comparison of single-pixel imaging algorithms," JOSA A 35, 78–87 (2018). [PubMed: 29328095]

15. Sampsell JB, "Digital Micromirror Device and Its Application to Projection Displays," J Vac Sci Technol B 12, 3242–3246 (1994).

16. Xu ZH, Chen W, Penuelas J, Padgett M, and Sun MJ, "1000 fps computational ghost imaging using LED-based structured illumination," Opt Express 26, 2427–2434 (2018). [PubMed: 29401782]

17. Wu H, Zheng Z, Li Y, Dai W, and Xiong H, "Compressed Sensing via a Deep Convolutional Auto-encoder," in 2018 IEEE Visual Communications and Image Processing (VCIP)(IEEE2018), pp. 1–4.

18. Zhang J, and Ghanem B, "ISTA-Net: Interpretable optimization-inspired deep network for image compressive sensing," in Proceedings of the IEEE conference on computer vision and pattern recognition(2018), pp. 1828–1837.

19. Beck A, and Teboulle M, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," SIAM journal on imaging sciences 2, 183–202 (2009).

20. Pavillon N, and Smith NI, "Compressed sensing laser scanning microscopy," Opt Express 24, 30038–30052 (2016). [PubMed: 28059389]

21. Fang L, Monroe F, Novak SW, Kirk L, Schiavon CR, Seungyoon BY, Zhang T, Wu M, Kastner K, and Kubota Y, "Deep learning-based point-scanning super-resolution imaging," bioRxiv, 740548 (2019).

22. Li B, Yan Q-R, Wang Y-F, Yang Y-B, and Wang Y-H, "A binary sampling Res2net reconstruction network for single-pixel imaging," Review of Scientific Instruments 91, 033709 (2020).

23. Wang H, Rivenson Y, Jin Y, Wei Z, Gao R, Gunaydin H, Bentolila LA, Kural C, and Ozcan A, "Deep learning enables cross-modality super-resolution in fluorescence microscopy," Nat Methods 16, 103–110 (2019). [PubMed: 30559434]

24. Deng J, Dong W, Socher R, Li L-J, Li K, and Fei-Fei L, "Imagenet: A large-scale hierarchical image database," in 2009 IEEE conference on computer vision and pattern recognition(Ieee2009), pp. 248–255.

25. Kulkarni K, Lohit S, Turaga P, Kerviche R, and Ashok A, "Reconnet: Non-iterative reconstruction of images from compressively sensed measurements," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(2016), pp. 449–458.

26. Martin D, Fowlkes C, Tal D, and Malik J, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001(IEEE2001), pp. 416–423.

27. Koho S, Tortarolo G, Castello M, Deguchi T, Diaspro A, and Vicidomini G, "Fourier ring correlation simplifies image restoration in fluorescence microscopy," Nature communications 10, 1–9 (2019).

28. Lehmann TM, Gonner C, and Spitzer K, "Addendum: B-spline interpolation in medical image processing," IEEE Transactions on Medical Imaging 20, 660–665 (2001). [PubMed: 11465471]

29. Pologruto TA, Sabatini BL, and Svoboda K, "ScanImage: flexible software for operating laser scanning microscopes," Biomedical engineering online 2, 13 (2003). [PubMed: 12801419]

30. Ronneberger O, Fischer P, and Brox T, "U-net: Convolutional networks for biomedical image segmentation," in International Conference on Medical image computing and computer-assisted intervention(Springer2015), pp. 234–241.

31. Feng J, Deng J, Li Z, Sun Z, Dou H, and Jia K, "End-to-end Res-Unet based reconstruction algorithm for photoacoustic imaging," Biomedical Optics Express 11, 5321–5340 (2020). [PubMed: 33014617]

32. Chen Y, Xie Y, Zhou Z, Shi F, Christodoulou AG, and Li D, "Brain MRI super resolution using 3D deep densely connected neural networks," in 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)(IEEE2018), pp. 739–742.

33. Huang G, Liu Z, Van Der Maaten L, and Weinberger KQ, "Densely connected convolutional networks," in Proceedings of the IEEE conference on computer vision and pattern recognition(2017), pp. 4700–4708.

34. Lee B, Ku B, Kim W-J, Kim S, and Ko H, "Denoising ISTA-Net: learning based compressive sensing with reinforced non-linearity for side scan sonar image denoising," The Journal of the Acoustical Society of Korea 39, 246–254 (2020).

35. Nicola Pezzotti E. d. W., Yousefi Sahar, Elmahdy Mohamed S., van Gemert Jeroen, Schülke Christophe, Doneva Mariya, Nielsen Tim, Kastryulin Sergey, Lelieveldt Boudewijn P.F., van Osch Matthias J.P., Staring Marius, "Adaptive-CS-Net: FastMRI with Adaptive Intelligence," arXiv: 1912.12259 (2019).

36. Zhou XQ, Hou YH, and Lin JQ, "A review on the processing accuracy of two-photon polymerization," Aip Advances 5 (2015).

37. Gissibl T, Thiele S, Herkommer A, and Giessen H, "Two-photon direct laser writing of ultracompact multi-lens objectives," Nature Photonics 10, 554–+ (2016).

38. Li JW, Fejes P, Lorenser D, Quirk BC, Noble PB, Kirk RW, Orth A, Wood FM, Gibson BC, Sampson DD, and McLaughlin RA, "Two-photon polymerisation 3D printed freeform micro-optics for optical coherence tomography fibre probes," Sci Rep-Uk 8 (2018).

39. Mohammad N, Meem M, Wan X, and Menon R, "Full-color, large area, transmissive holograms enabled by multi-level diffractive optics," Sci Rep 7, 5789 (2017). [PubMed: 28724934]

40. Vaidya N, and Solgaard O, "3D printed optics with nanometer scale surface roughness," Microsystems & Nanoengineering 4, 1–8 (2018). [PubMed: 31057891]

41. Wu YC, Boominathan V, Chen HJ, Sankaranarayanan A, and Veeraraghavan A, "PhaseCam3D-Learning Phase Masks for Passive Single View Depth Estimation," Ieee Int Conf Comput (2019).

42. Sitzmann V, Diamond S, Peng YF, Dun X, Boyd S, Heidrich W, Heide F, and Wetzstein G, "End-to-end Optimization of Optics and Image Processing for Achromatic Extended Depth of Field and Super-resolution Imaging," Acm T Graphic 37 (2018).

43. Sun QL, Zhang J, Dun X, Ghanem B, Peng YF, and Heidrich W, "End-to-end Learned, Optically Coded Super-resolution SPAD Camera," Acm T Graphic 39 (2020).

44. Li Y, Sixou B, and Peyrin F, "A review of the deep learning methods for medical images super resolution problems," IRBM (2020).

45. Ouyang W, Aristov A, Lelek M, Hao X, and Zimmer C, "Deep learning massively accelerates super-resolution localization microscopy," Nat Biotechnol 36, 460–468 (2018). [PubMed: 29658943]

46. Weigert M, Schmidt U, Boothe T, Muller A, Dibrov A, Jain A, Wilhelm B, Schmidt D, Broaddus C, Culley S, Rocha-Martins M, Segovia-Miranda F, Norden C, Henriques R, Zerial M, Solimena M, Rink J, Tomancak P, Royer L, Jug F, and Myers EW, "Content-aware image restoration: pushing the limits of fluorescence microscopy," Nat Methods 15, 1090–1097 (2018). [PubMed: 30478326]

47. Zhang S, "High-speed 3D shape measurement with structured light methods: A review," Optics and Lasers in Engineering 106, 119–131 (2018).

48. Wu YC, Rivenson Y, Wang HD, Luo YL, Ben-David E, Bentolila LA, Pritz C, and Ozcan A, "Three-dimensional virtual refocusing of fluorescence microscopy images using deep learning," Nat Methods 16, 1323–+ (2019). [PubMed: 31686039]

49. Pinkard H, Phillips Z, Babakhani A, Fletcher DA, and Waller L, "Deep learning for single-shot autofocus microscopy," Optica 6, 794–797 (2019).

50. Ma Y, Lee Y, Best-Popescu C, and Gao L, "High-speed compressed-sensing fluorescence lifetime imaging microscopy of live cells," bioRxiv (2020).

51. Qi D, Zhang S, Yang C, He Y, Cao F, Yao J, Ding P, Gao L, Jia T, and Liang J, "Single-shot compressed ultrafast photography: a review," Advanced Photonics 2, 014003 (2020).

52. Gao L, Liang JY, Li CY, and Wang LHV, "Single-shot compressed ultrafast photography at one hundred billion frames per second," Nature 516, 74–U159 (2014). [PubMed: 25471883]

53. Alemohammad M, Shin J, Tran DN, Stroud JR, Chin SP, Tran TD, and Foster MA, "Widefield compressive multiphoton microscopy," Optics letters 43, 2989–2992 (2018). [PubMed: 29905741]

54. Wijesinghe P, Escobet-Montalban A, Chen MZ, Munro PRT, and Dholakia K, "Optimal compressive multiphoton imaging at depth using single-pixel detection," Opt Lett 44, 4981–4984 (2019). [PubMed: 31613244]

55. Ren Y-X, Kong C, He H, Zeng X, Tsia KK, and Wong KK, "Encrypted wide-field two-photon microscopy with single-pixel detection and compressed sensing," Applied Physics Express 13, 032007 (2020).

56. Stirman JN, Smith IT, Kudenov MW, and Smith SL, "Wide field-of-view, multi-region, two-photon imaging of neuronal activity in the mammalian brain," Nature biotechnology 34, 857–862 (2016).
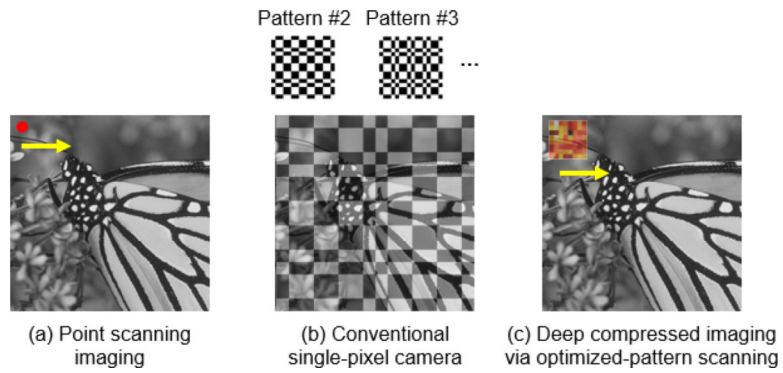
Fig. 1.

The landscape of imaging methods using a single-pixel detector. (a) Point scanning system where signal from individual pixel is sequentially recorded. (b) A conventional single-pixel camera where different patterns are sequentially projected on the entire object and the overlap integrals between the object and each pattern are measured. (c) Deep compressed imaging via optimized-pattern scanning (DeCIOPS), where a pattern is scanned across the object and the subsampled convolution between the pattern and the object is measured.
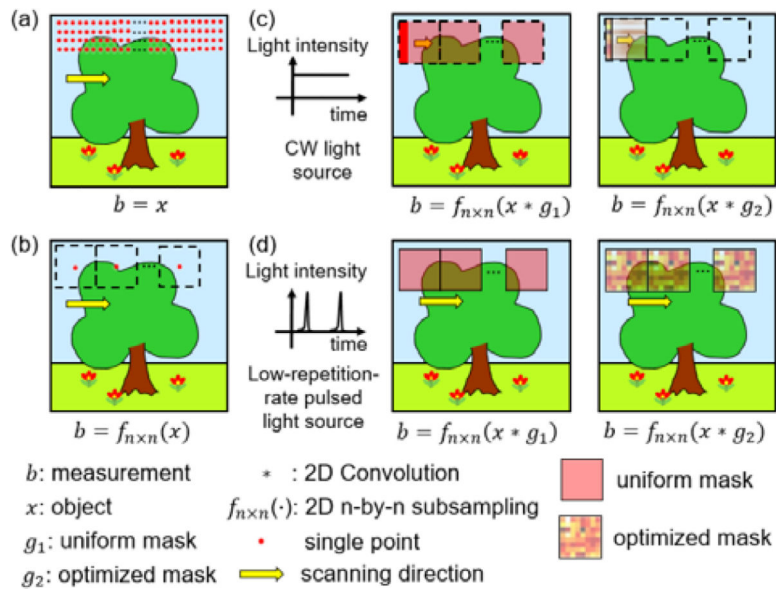
Fig. 2.
Schematic of the undersampling schemes in DeCIOPS. (a) Conventional pixel-by-pixel point-scanning. (b) Pixel-by-pixel point-scanning with a simple undersampling scheme. (c) DeCIOPS in a CW light source configuration with an illumination pattern of a uniform mask (left) or an optimized mask (right). (d) DeCIOPS in a low-repetition-rate pulsed light source configuration with a uniform mask (left) or an optimized mask (right) as an illumination pattern. The mathematic formula below each panel illustrates the process of image formation, where g_1 and g_2 are both in square shape.
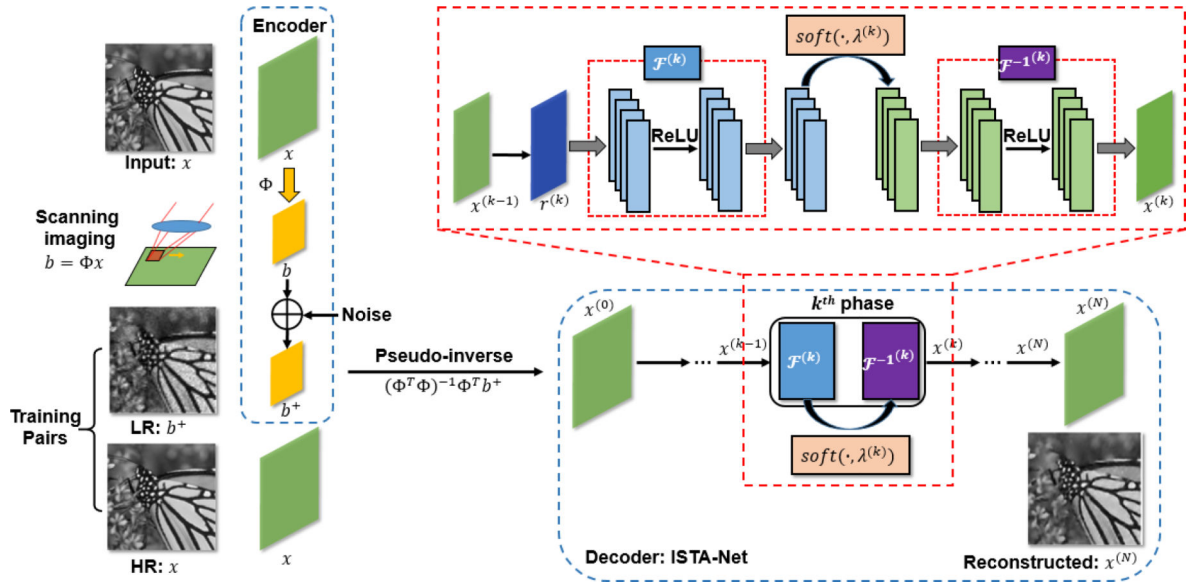
Fig. 3.

An end-to-end optimized auto-encoder framework of image formation and reconstruction in DeCIOPS. The encoder models the image formation. It encodes the high resolution (HR) object x into a low resolution (LR) output b^+ through subsampled convolution Φ and additive noise. The decoder is implemented with an ISTA-Net, which contains N phases and reconstructs the object $x^{(N)}$. Each phase is realized by a structure-symmetric pair of a forward transform $\mathscr{F}^{(k)}$ and a backward transform $\mathscr{F}^{-1(k)}$ with a soft shrinkage threshold, which factually matches one iteration in the conventional ISTA. ReLU, rectified linear unit. *Soft*($\cdot$): soft shrinkage threshold.
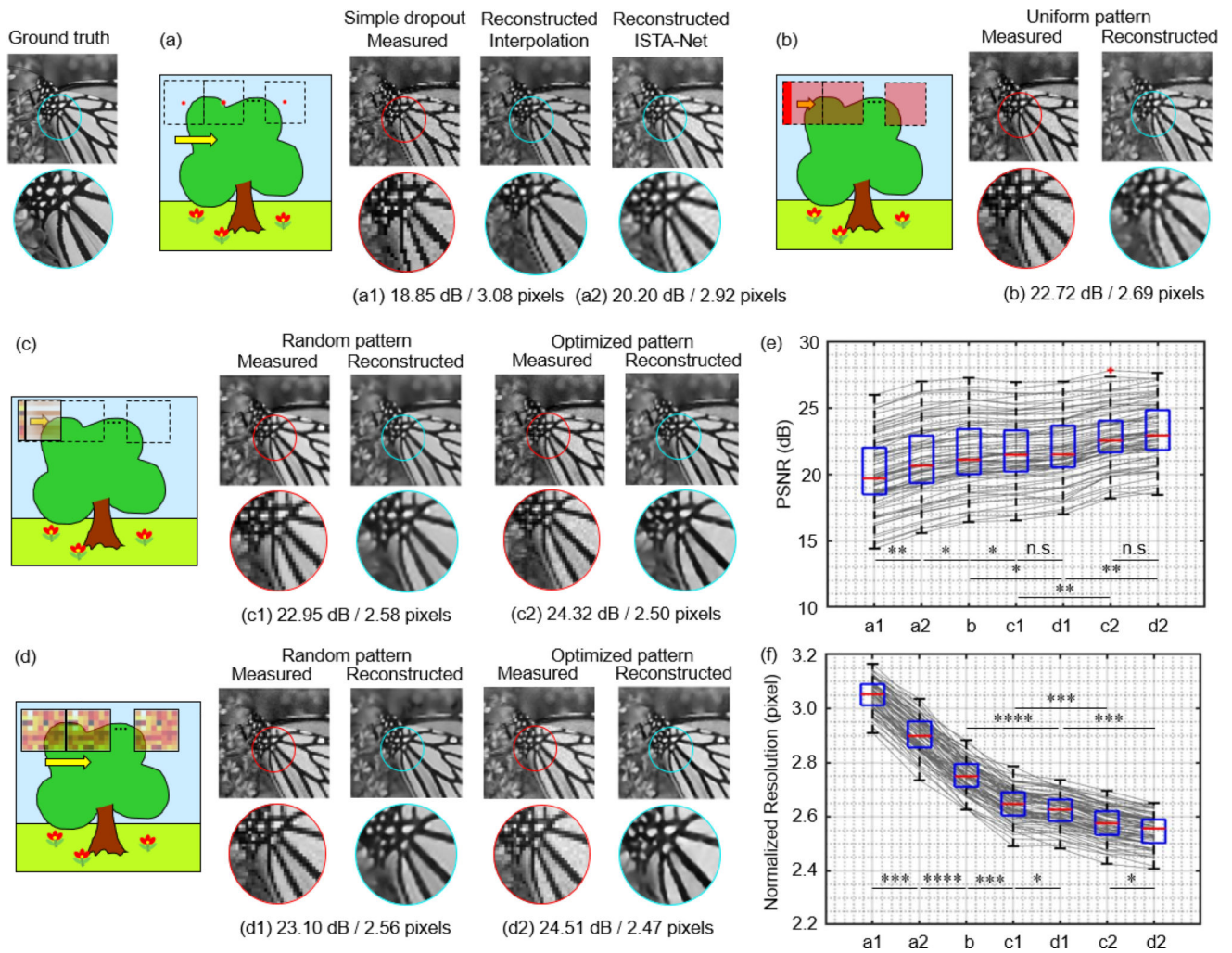
Fig. 4.

Comparison of the reconstruction performance in the validation data set Set11 and BSD68, at an undersampling rate of 6.25%, through (a) a simple dropout, (b) an unweighted average (uniform pattern), (c) a random or an optimized illumination pattern (DeCIOPS) with a constraint of identical column, and (d) a random or an optimized illumination pattern (DeCIOPS). The PSNR and resolution of the reconstructed images are labeled below the exemplary sample. (e) PSNR of the reconstructed images of all 79 samples in the validation dataset for cases in (a-d). (f) Resolution of the reconstructed images of all 79 samples in the validation dataset for cases in (a-d). n.s., not significant; *, $p<0.05$; **, $p<0.01$; ***, $p<0.001$; ****, $p<0.0001$, in one-way Analysis of Variance (ANOVA).
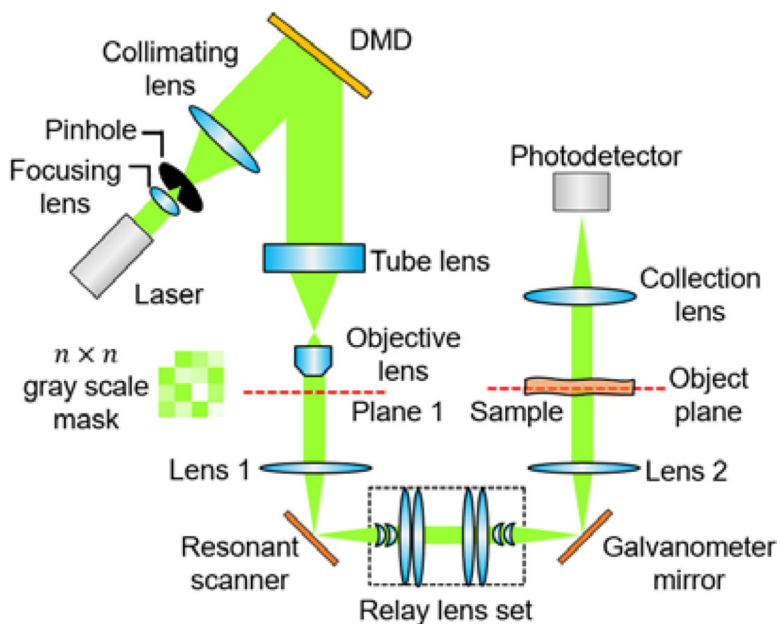
Fig 5.

The experiment setup of DeCIOPS. The laser beam is spatially filtered to improve its spatial uniformity and symmetricity, collimated and expanded in size, and then incident onto a DMD. The beam is spatially modulated by the DMD and then shrunk in size by a 4f system formed by a tube lens and an objective lens. The light pattern is scanned by a resonant-galvo scanner set, where a resonant scanner and a galvanometer mirror is optically coupled through a relay lens set. The transmitted light from the sample is collected by a photodetector through a collection lens. The $n \times n$ pattern is generated by the DMD. With an additional 4f system with cylindrical lenses after the objective lens, the $n \times n$ pattern can be turned into $n \times 1$ size (Appendix 1.A). The red dashed line (plane 1 and object plane) indicates the conjugate plane of the gray-scale pattern mask.
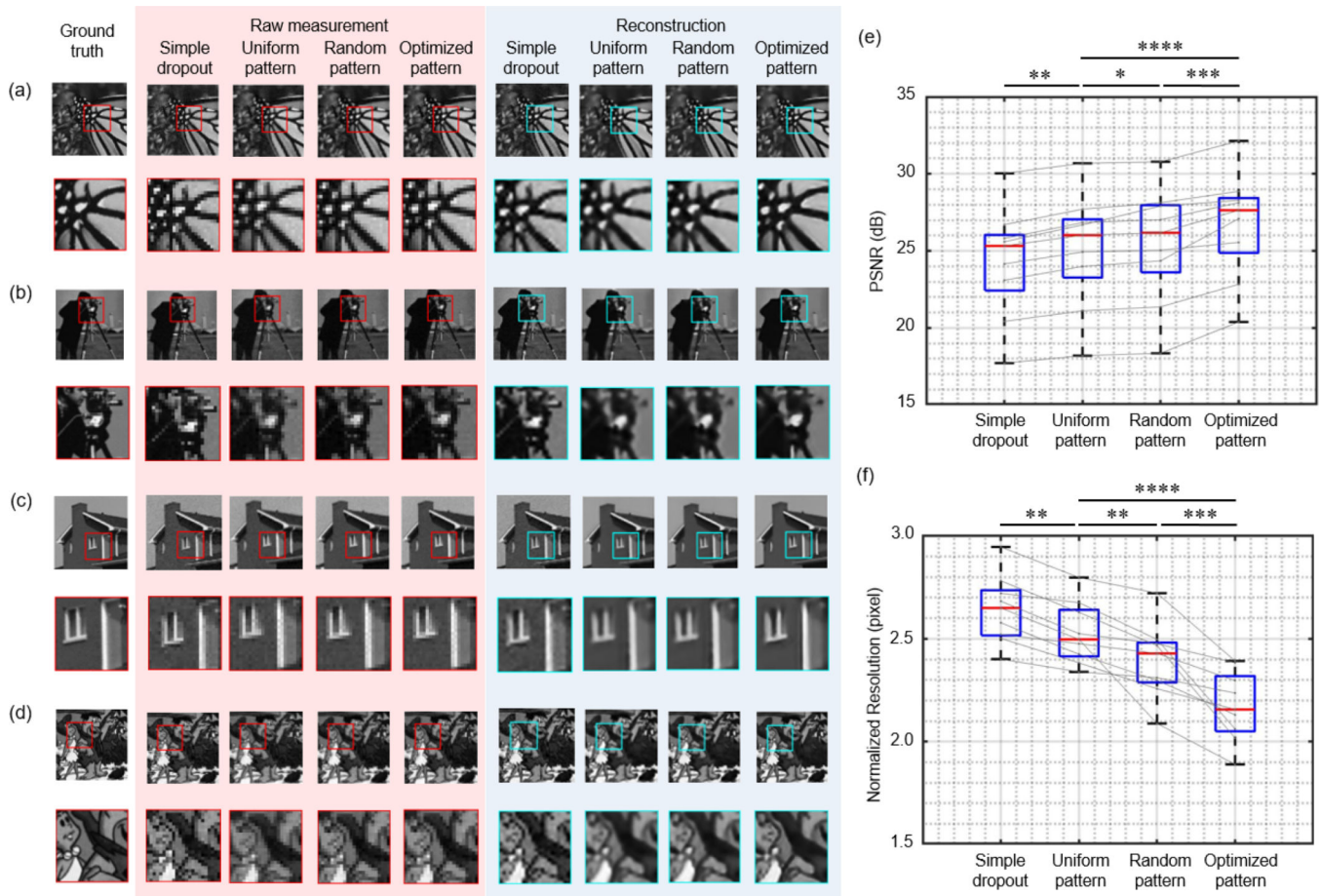
Fig. 6.

Comparison of the experimental results using different illumination patterns in the scanning in a CW illumination setting. (a-d), Experimental results of the sample (a) Butterfly, (b) Cameraman, (c) House and (d) Flinstones. The different columns show the ground truth results using high-resolution point-scanning, raw measurement using different illumination patterns at an undersampling rate of 6.25%, and the corresponding reconstruction results. (e) PSNR of the reconstructed images for a total of 9 samples. (f) Spatial resolution of the reconstructed images for a total of 9 samples, calculated from Fourier ring correlation. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$, in one-way Analysis of Variance (ANOVA).
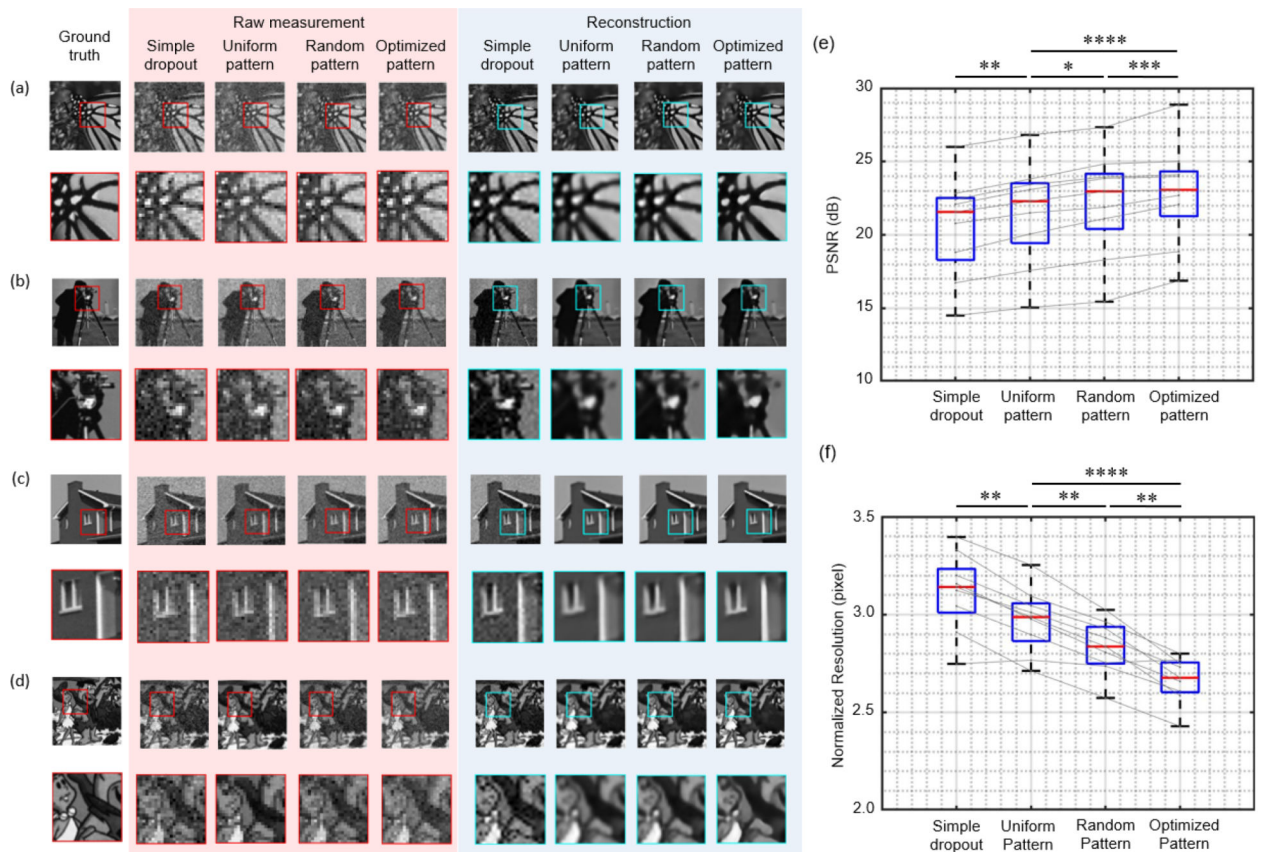
Fig. 7.

Comparison of the experimental results using different illumination patterns in the scanning in the low-repetition-rate pulsed light illumination setting. (a-d), Experimental results of the sample (a) Butterfly, (b) Cameraman, (c) House and (d) Flinstones. The different columns show the ground truth results using high-resolution point-scanning, raw measurement using different illumination patterns at an undersampling rate of 6.25%, and the corresponding reconstruction results. (e) PSNR of the reconstructed images for a total of 9 samples. (f) Spatial resolution of the reconstructed images for a total of 9 samples, calculated from Fourier ring correlation. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$; ****, $p < 0.0001$, in one-way Analysis of Variance (ANOVA).
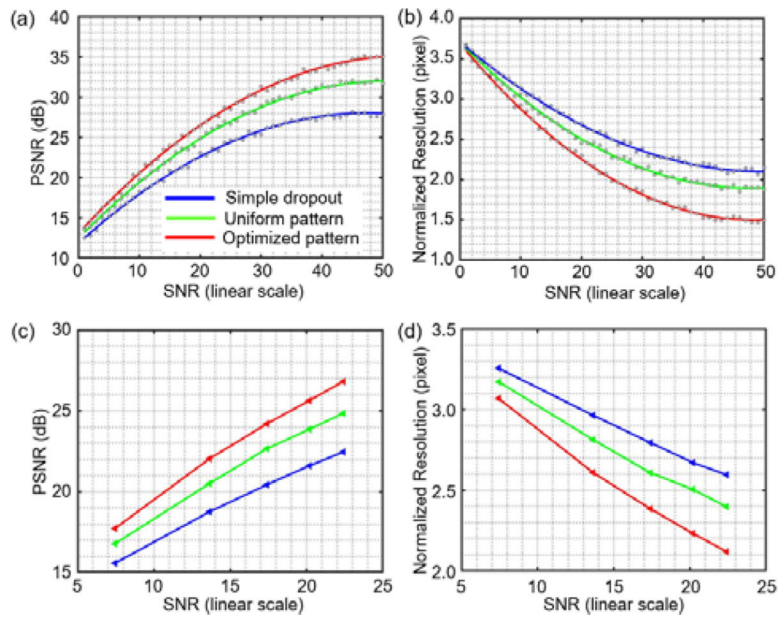
Fig. 8.
(a) PSNR and (b) pixel resolution of the reconstructed images versus different SNR in the raw measurement, for three different sampling patterns (CW configuration), performed through simulation, at an undersampling rate of 6.25%. The results were averaged across 9 samples used in the experiment and fitted with polynomial curves. (c) and (d) shows the experimental results, averaged across 9 samples.
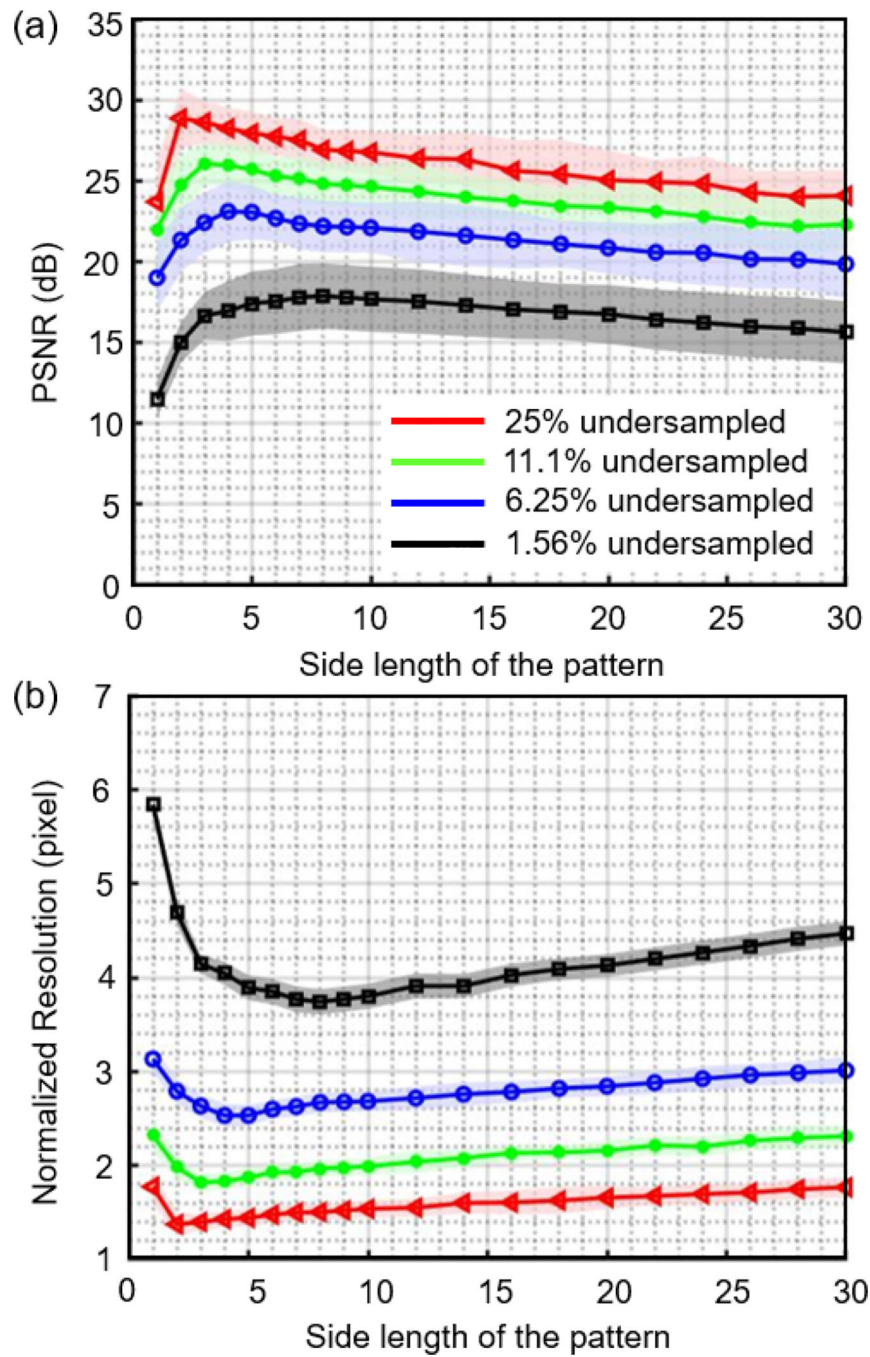
Fig. 9.
DeCIOPS reconstruction quality (a) PSNR and (b) pixel resolution dependence on the size of the optimized pattern, for an undersampling rate of 25% (2 ×2, red), 11.1% (3 ×3, green), 6.25% (4 × 4, blue) and 1.5625% (8 ×8, black), across all 79 samples in the validation dataset. Solid curve: mean; shaded area: standard deviation.
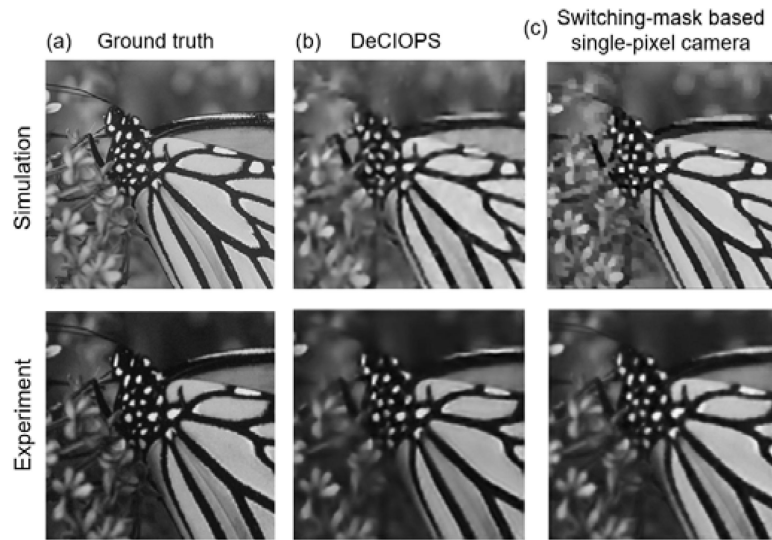
Fig. 10.

Comparison of the reconstruction results between DeCIOPS and conventional switching-mask based single-pixel camera. (a) The ground truth of an original object, butterfly. (b) Reconstruction result of DeCIOPS using ISTA-Net at an undersampling rate of 6.25%. (c) Reconstruction result of the switching-mask based single-pixel camera imaging approach using ISTA-Net. Top row, simulation. Bottom row, experiment. The ground truth of the experiment is obtained by the high-resolution point-scanning.

Table 1.

Comparison of the PSNR and pixel resolution across uniform pattern, random pattern and optimized pattern between 6.25% and 1.5625% undersampling rate in DeCIOPS (CW light configuration, SNR~22, for all 79 samples in validation dataset)

Performance metric	Undersampling rate	Uniform pattern	Random pattern	Optimized pattern
PSNR	6.25%	21.49±1.83 dB	21.82±1.72 dB	23.13±1.54 dB
	1.5625%	15.91±2.01 dB	16.78±1.94 dB	18.47±1.75 dB
Normalized resolution	6.25%	2.79±0.18 pixel	2.68±0.16 pixel	2.55±0.15 pixel
	1.5625%	4.55±0.26 pixel	4.28±0.25 pixel	3.76±0.21 pixel

Table 2.

Comparison of reconstruction results between DeCIOPS (CW light configuration) and conventional switching-mask based single-pixel camera. Both simulation and experiment were performed on 9 samples from the validation dataset, with an undersampling rate of 6.25% and SNR∼22. No significant difference (one-way ANOVA) is found on the performance metrics between the two imaging modalities.

Performance metric	Simulation		Experiment	
	PSNR	**Normalized Resolution**	**PSNR**	**Normalized Resolution**
DeCIOPS	28.01±1.03 dB	2.13±0.12 pixel	27.71±1.18 dB	2.15±0.13 pixel
Switching-mask based single-pixel camera	28.36±1.00 dB	2.12±0.12 pixel	27.98±1.16 dB	2.14±0.13 pixel

Table 3.

Detailed parameters of the optical components used in the imaging system.

Element	Manufacturer	Part Number	Note
Focusing lens	Thorlabs	A397TM-A	Aspherical lens, focal length 11 mm
Pinhole	Thorlabs	P30D	30 μm pinhole
Collimating lens	Thorlabs	AC254–100-AB-ML	Achromatic lens, focal length 100 mm (for pattern scanning)
	Thorlabs	ACL5040-A	Aspherical lens, focal length 40 mm (for point scanning)
Tube lens	Thorlabs	SM2V10	Focal length 200 mm
Objective lens	Olympus	RMS4X	4 × objective lens
Cylindrical lens	Throlabs	LJ1014L1-A	Focal length 25.4 mm
	Throlabs	LJ1227L2-A	Focal length 6.35 mm
Lens 1 and 2	OptoSigma	SLB-60–250P	Focal length 250 mm
Relay lens set [56] (one set listed here; two sets arranged symmetrically are required in setup)	Throlabs	LC1582-A	Focal length -75 mm
	Thorlabs	LC1582-A	Focal length -75 mm
	Thorlabs	LE1076-A	Meniscus Lens, focal length 100 mm
	Thorlabs	LA1399-A	Focal length 175 mm
	Thorlabs	LA1050-A	Focal length 100 mm
	Thorlabs	LA1727-A	Focal length 750 mm
Collection lens	Thorlabs	AC508–075-A-ML	Achromatic lens, focal length 75 mm
Galvometric scanner	Cambridge Technology	6215HM40B	
Resonant scanner	Cambridge Technology	CRS 8 KHz	Resonant frequency 8 KHz
DMD	Texas Instruments	DLP7000	