

Resolution enhancement of long-range imaging with sparse apertures

Jiachen Wu, Feng Yang, Liangcai Cao*

State Key Laboratory of Precision Measurement Technology and Instruments, Department of Precision Instruments, Tsinghua University, Beijing, 100084, China

ARTICLE INFO

Keywords:

Fourier ptychography
Synthetic aperture
Total variation regularization

ABSTRACT

Synthetic aperture technique could effectively expand the passband and beyond the diffraction limit of aperture. Macroscopic Fourier ptychography (FP) makes synthetic aperture feasible in visible light by incorporating phase retrieval process into synthetic aperture. In FP, a dense aperture sampling is necessary to overcome the ill-posedness of phase retrieval, which leads to long acquisition time and heavy computation. Here, a total variation regularization-based Fourier ptychographic (TVFP) reconstruction method is proposed for long-range imaging. The TVFP method imposes total variation constraints to improve the ill-posedness of the imaging model, and recovers high-quality image from sparse aperture sampling. For the imaging system with aberration, a joint optimization of pupil function and target image is also proposed to achieve the aberration free image and the estimated pupil function simultaneously. The results show the aperture overlapping ratio could be reduced to 25% or less. The proposed method simplifies the imaging configuration and improves the robustness for long-range imaging.

1. Introduction

High resolution image acquisition at long range is challenging in the surveillance and remote sensing fields. For a diffraction-limited system, when the distant light wave arrives at the aperture plane, the finite size of the aperture disturbs the propagation of the wavefront, and leads to diffraction blurring. According to the Rayleigh criterion, the minimum resolvable distance is inversely proportional to the diameter of the aperture. Increasing the size of the aperture of the lens is a straightforward way to improve the resolution. However, a large aperture results in high expense and bulky volume, which is unfeasible in most scenarios. Therefore, the modality of synthetic aperture is extensively applied in long-range imaging [1–3].

Synthetic aperture imaging is a mature technology in long-wavelength regimes, such as synthetic aperture radar (SAR) [4] and very-long-baseline interferometry (VLBI) [5]. The azimuth resolution of the radar is related to the antenna aperture size. SAR is typically mounted on a moving platform, such as an aircraft or spacecraft. When the antenna moves relative to the target, an equivalent large antenna aperture can be obtained in the direction of motion. In VLBI, multiple radio telescopes are combined to observe an object simultaneously. VLBI can be used to synthesize high-resolution images of radio sources by measuring their complex visibility. In these methods, both the intensity and phase of the incident field can be directly measured. For visible wavelengths, imaging sensors could only record intensity, and all phase information was lost, which

means these methods could not be equivalently effective for visible light.

The Fourier ptychography (FP) technique, which combines phase retrieval process with synthetic aperture technique, has been applied in microscopic [6–8] and macroscopic imaging [9–11]. This technique uses one or multiple low numerical aperture (NA) optical systems to capture a set of low-resolution images that correspond to different parts of the Fourier spectrum, and synthesizes these measurements to expand the passband and improve the achievable resolution. Because the sensor can only measure the magnitude of the signal, phase retrieval algorithms need to be incorporated into FP reconstruction. Alternating projection is a kind of widely used phase retrieval framework. [6,12,13]. More advanced optimization strategies also have FP implementations such as Wirtinger flow [14,15], convex relaxation [16] and low-rank recovery [17]. However, phase retrieval problem is ill-posed due to the inherent ambiguity of solution, such as piston diversity [18] and twin-image problem [19]. The overlapping between each measurement in the spectrum domain is necessary in current FP algorithms to ensure the consistency of the solutions. A high overlapping ratio provides a fast convergence rate, but also results in a long time for data acquisition and computation. Bunk et al. suggested 30% overlapping for high speed ptychographic measurement [20]. Dong et al. investigated the spectral-spatial data redundancy and showed that a minimum of ~35% overlapping is needed for successful FP reconstruction [21]. Indeed, the ambiguity of the phase retrieval problem could also be addressed by introducing regularization techniques that impose certain prior distri-

* Corresponding author at: Dr Liangcai Cao, Tsinghua University, Room 3102, Building 9003, Beijing 100084, China.

E-mail address: clc@tsinghua.edu.cn (L. Cao).

butions on model parameters [22–24]. Although several regularization models have been built for FP reconstruction [25–28], these methods aim to depress the noise, but the prior knowledge of image is not fully exploited. Jagatap et al. proposed a subsampling strategy by utilizing the underlying structured sparsity in an object, which is limited to the reconstruction of block-sparse objects [29]. The image priors could also be learned from large amounts of data using deep learning methods [30–33]. Because deep learning is a data-driven method, the performance is mainly affected by the correlation between measurements and training data.

The total variation (TV) regularization was introduced by Rudin et al. [34] for image denoising, and then evolved into a more general technique for inverse problems, including FP reconstruction for microscope [35,36]. The principle of TV regularization is based on the image prior, assuming that the gradient of the natural scene is sparse, in other words that the image is piecewise smooth in appearance. For a problem with no unique solution, or where the solution is sensitive to noise, reducing the TV of the target image tends to seek a solution that is a close match to the original image. The noise, artifacts, and any spurious details causing high TV value are eliminated. The excellent properties of TV regularization are well suited for reducing the redundant FP measurements.

In this work, a total variation regularization-based Fourier ptychography (TVFP) method is proposed for long-range imaging with sparse aperture, aiming to achieve high resolution imaging with fewer measurements. The TV regularization strengthens the phase consistency in the spectral domain of each measurement to overcome the ill-posedness, and provides a reasonable guess of non-captured spatial frequencies. To incorporate the sparse prior into the phase retrieval problem, the alternating direction method (ADM) is elaborated for FP reconstruction instead of conventional projection algorithms. Meanwhile, an extended TVFP (eTVFP) method is proposed to estimate the pupil function under a joint optimization. The estimation of pupil function allows the calibration of aberration to be included in the iterations to keep the fidelity of the imaging system. The TVFP method reduce the overlapping ratio to 25% or less for reconstruction of natural scene. The proposed methods can prompt the data processing efficiency of the FP to a new horizon.

2. Observation model

In the case of long-range imaging, when the light from an object propagates over a sufficiently large distance z toward the camera, the light field at the aperture plane of the camera can be calculated with the far-field Fraunhofer approximation. The light field at the aperture plane is related to the field at the object through a Fourier transform. For the non-self-luminous object, a coherent light beam is used to illuminate the distant object, and the object reflects the light toward the camera. The optical setup is shown in Fig. 1. The focusing lens expands the laser beam and focuses light source on the aperture plane. According to the Fresnel diffraction formula [37], the field H at the aperture plane of the camera can be represented as:

$$H(u, v) = \frac{e^{jkz}}{j\lambda z} \exp \left[j \frac{k}{2z} (u^2 + v^2) \right] \mathcal{F} \{ O(x_o, y_o) \} \Big|_{\xi=\frac{u}{\lambda z}, \eta=\frac{v}{\lambda z}} \quad (1)$$

where $k = 2\pi/\lambda$ is the wave number, λ is the illumination wavelength, O is the light field at the object plane, z is the propagation distance from the object plane to the aperture plane. ξ and η are the spatial frequency, u and v are physical spatial coordinates in aperture plane. Eq. (1) shows that the field H at the aperture plane differs from the spectrum of the object by only a quadratic phase factor.

When the field H continues to propagate forward, the optical system blocks the wave outside the aperture, and only a portion of H can enter the interior of the optical system, which can be expressed through multiplication by a pupil function. For an aberration-free optical system, the

pupil function is defined as:

$$P(u, v) = \begin{cases} 1, & \text{if } \sqrt{u^2 + v^2} < D/2 \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

where D is the diameter of the pupil. The optical field arriving at the sensor plane is obtained by using Fresnel diffraction formula again:

$$\begin{aligned} U(x, y) &= \frac{e^{jkd}}{j\lambda d} \exp \left[j \frac{k}{2d} (x^2 + y^2) \right] \\ &\times \mathcal{F} \left\{ \exp \left[j \frac{k}{2} (u^2 + v^2) \right] \left(\frac{1}{d} - \frac{1}{f} \right) \right\} H(u, v) P(u, v) \Big\} \\ &= -\frac{e^{jk(d+z)}}{\lambda^2 d z} \exp \left[j \frac{k}{2d} (x^2 + y^2) \right] \\ &\times \mathcal{F} \{ \psi(u, v) P(u, v) \} \Big|_{\xi=\frac{x}{\lambda z}, \eta=\frac{y}{\lambda z}} \end{aligned} \quad (3)$$

where d is the image distance, f is the focal length of imaging lens, and $\psi = \mathcal{F}\{O\}$ is the spectrum of the object. Eq. (3) holds when $1/f = 1/d + 1/z$, which indicates that the sensor plane and the object plane are conjugated. For simplicity, we use vectors \mathbf{x} and \mathbf{u} to represent the object coordinates (x, y) and aperture coordinates (u, v) respectively. Because the image of object cast on the sensor is inverted, the inverse Fourier transform operator \mathcal{F}^{-1} is used instead of \mathcal{F} to invert the sensor plane coordinates. The intensity captured by the sensor is:

$$I(\mathbf{x}, \mathbf{c}) = |U(\mathbf{x})|^2 \propto \left| \mathcal{F}^{-1} \{ \psi(\mathbf{u} - \mathbf{c}) P(\mathbf{u}) \} \right|^2 \quad (4)$$

where \mathbf{c} is the offset vector determined by the camera position or the angle of illumination. Recovering the high-resolution image of an object is equivalent to recovering the complete spectrum of an object, which can be achieved through aperture scanning. Assuming that the number of apertures or measurement times is N , meaning that there are N different offset vectors \mathbf{c}_i , $i = 1, 2, \dots, N$, the complete relative spectrum distribution can be estimated by solving the optimization problem:

$$\min_{\psi} \sum_{i=1}^N \left\| \mathcal{F}^{-1} \{ \psi(\mathbf{u} - \mathbf{c}_i) P(\mathbf{u}) \} - b_i \right\|_2^2 \quad (5)$$

where b_i is the measured amplitude, which is the square root of the intensity I_i . There are two key challenges to solve Eq. (5). The first is that the phase of the measurements is lost because the sensor can only capture the intensity. The second is that the apertures cannot thoroughly cover the spectrum plane when the number of measurements is insufficient, which means that the sampling of the Fourier spectrum is incomplete. Most FP algorithms only solve the first issue by performing phase retrieval for each aperture. If complete spectrum information is desired, the only way is to increase the number of measurements until the measurements cover the entire spectrum. In addition, the overlapping of the apertures is a necessary condition for FP reconstruction. If there is no overlapping between apertures, FP reduces to a conventional phase retrieval procedure where each measurement is processed independently. The ambiguity of phase retrieval then makes each retrieved phase inconsistent, and the solution cannot converge to a desired value. In this work, a regularization term $\Phi(\psi)$ can be introduced into Eq. (5) as follows:

$$\min_{\psi} \tau \Phi(\psi) + \sum_{i=1}^N \left\| \mathcal{F}^{-1} \{ \psi(\mathbf{u} - \mathbf{c}_i) P(\mathbf{u}) \} - b_i \right\|_2^2 \quad (6)$$

where τ is a coefficient that balances the regularization term and the data fitting term. The regularization term will impose a prior distribution on the spectrum. The prior distribution provides extra information for the phase retrieval, which makes the convergence less dependent on the spectrum overlapping ratio. In the proposed TVFP method, $\Phi(\psi)$ is TV regularization term. We validate a minimum of 25% overlapping can ensure convergence for complex object. It can be even less in sparse amplitude-only cases. The prior distribution could also help to reasonable guess the spectrum outside the sampling region. So under the same

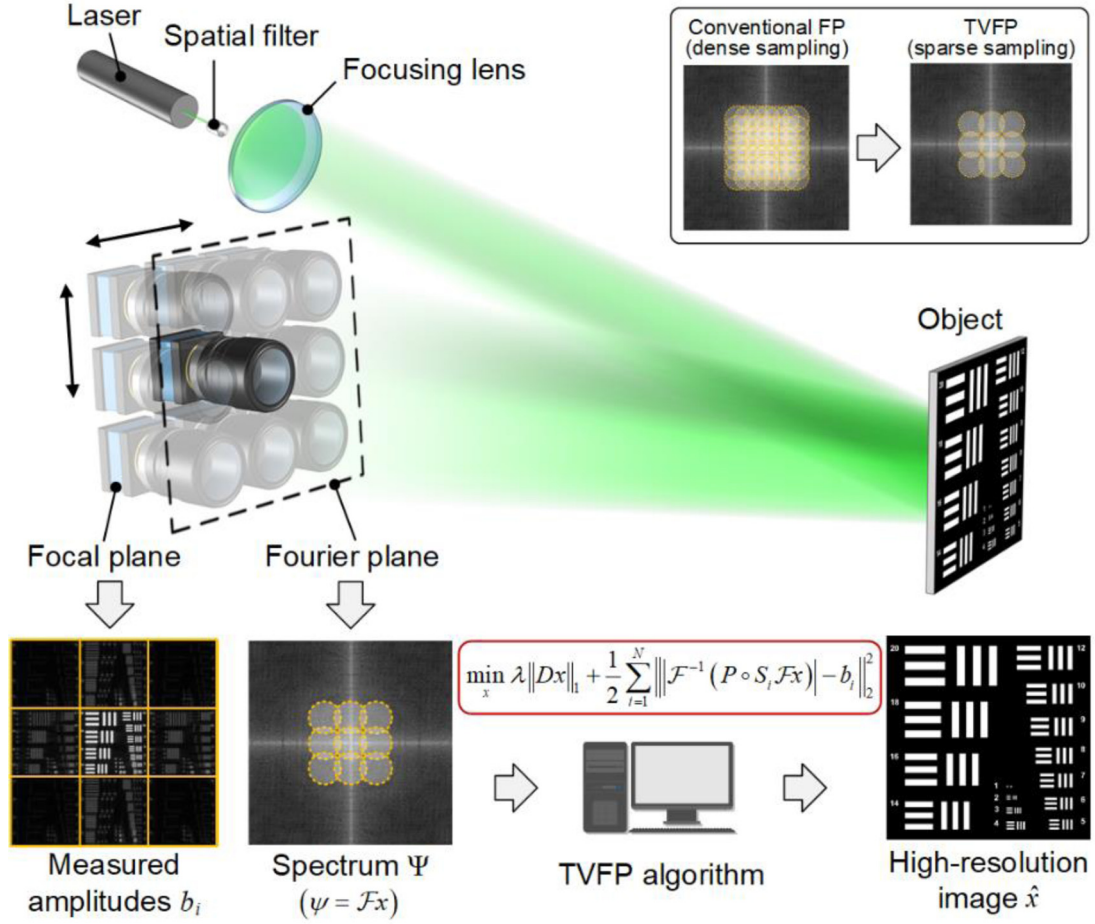


Fig. 1. An overview of proposed long-range imaging framework. The object is illuminated by a coherent beam, and the Fourier spectrum is formed at the aperture plane. With TVFP algorithm, dense sampling is reduced to sparse sampling as inset shows.

number of measurements, TVFP method can obtain much higher peak signal to noise ratio (PSNR) compared with the conventional FP algorithm.

3. Reconstruction with sparse aperture

Sparse aperture imaging is essentially an incomplete sampling in the spectral domain of target image. It is a special case of phase retrieval from incomplete spectrum sampling in which the sampling pattern is multiple circular apertures. In this section, phase retrieval from incomplete spectrum sampling via TV regularization is investigated, and then the approach can be easily applied to sparse aperture imaging. The derivation of the TV regularization method starts from a single measurement, and the excellent performance of TV regularization is shown more distinctly. Then the TV regularization method is generalized to multiple apertures to form TVFP method.

3.1. Phase retrieval from incomplete spectrum sampling

For a digital imaging sensor, the pixel pitch determines the maximum sampling frequency and delimits the available range of sampling in the spectral domain. The extent of spectrum sampling in the available range that would be sufficient for image recovery is of concern. Zhou et al. revealed that the result with half the number of captured data does not differ significantly from that with all the captured data when the sample is close to amplitude-only [38]. In compressed sensing (CS), the methods of signal recovery from incomplete spectrum data have been investigated and extensively applied in magnetic resonance

imaging (MRI) [39–41]. However, these works do not involve phase retrieval methods because the measured data are complex. Some methods of phase retrieval from incomplete amplitude information via ℓ_1 minimization [23] and TV regularization [24] have been proposed, but the measurements are in the Fourier domain. Inspired by previous works, we formulated a model of phase retrieval from incomplete spectrum sampling via TV regularization, which imposes a sparsity prior on the gradient domain of the underlying image. This overcomes the ill-posedness of phase retrieval and effectively improves the image quality.

Assuming that the spectrum of a signal x is sampled by a binary mask M , the regions corresponding to value ‘1’ in M denote the sampling regions in the Fourier domain. The amplitude b of the signal is then measured in the spatial domain. If we let $N = 1$, $P = M$, and $\psi = \mathcal{F}\{x\}$ in Eq. (6), and adopt TV regularization, the signal estimation can be posed in a linear operator form:

$$\min_x \tau \|Dx\|_1 + \frac{1}{2} \left\| \mathcal{F}^{-1}(M \circ \mathcal{F}x) - b \right\|_2^2 \quad (7)$$

where $D = [D_h, D_v]^T$ denotes the gradient operator, and D_h, D_v are the horizontal and vertical gradient operators, respectively. $\|Dx\|_1$ represents the anisotropic version of TV regularization term. Eq. (7) is non-convex owing to the presence of the absolute value operator. Here, we adopt the alternating direction method (ADM) [42], whose basic idea is to separate a complex problem into several subproblems that are easy to solve. The equivalent constrained problem of Eq. (7) is:

$$\begin{aligned} \min \tau \|w\|_1 + \frac{1}{2} \|g - b\|_2^2, \\ \text{s.t. } \mathcal{F}g = M \circ \psi, w = D\mathcal{F}^{-1}\psi \end{aligned} \quad (8)$$

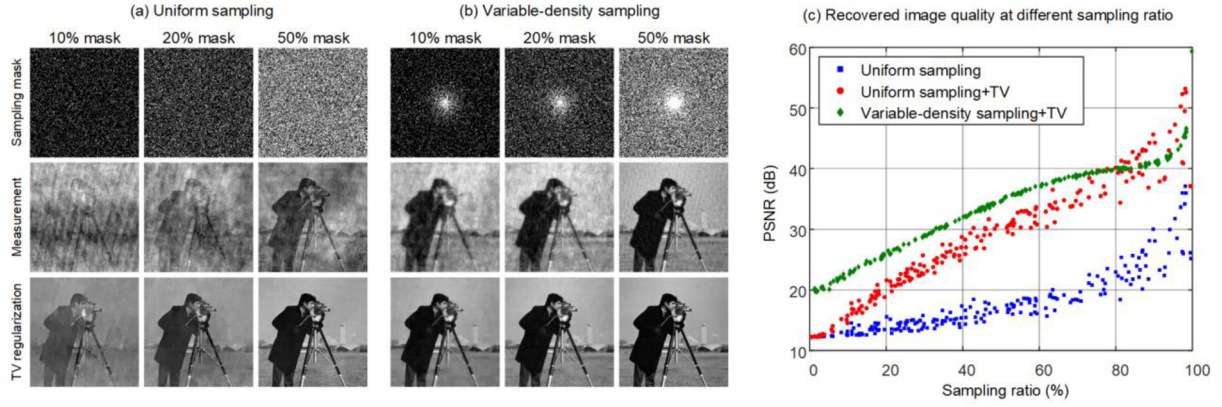


Fig. 2. Phase retrieval with different spectrum sampling ratios. Measured images and recovered images are shown with 10%, 20%, and 50% sampling ratio under (a) uniform sampling and (b) variable-density sampling, where $p = 20$. (c) The PSNR of recovered images are plotted at varying sampling ratios. Variable-density sampling with TV regularization provides better quality than uniform sampling.

The augmented Lagrangian function of the problem is given by:

$$\mathcal{L}(\psi, w, g, y, v) = \tau \|w\|_1 + \frac{1}{2} \|g - b\|_2^2 + \frac{\mu}{2} \|Fg - M \circ \psi + \frac{1}{\mu} y\|_2^2 + \frac{\eta}{2} \|w - DF^{-1}\psi + \frac{1}{\eta} v\|_2^2 \quad (9)$$

where y and v are the Lagrangian multipliers, and μ and η are the penalty parameters. Starting with ψ^0 , w^0 , y^0 , and v^0 , the ADM solves the following three subproblems sequentially in each iteration:

$$\begin{cases} g^{k+1} = \arg \min_{\psi} \mathcal{L}(\psi^k, w^k, g^k, y^k, v^k) \\ \psi^{k+1} = \arg \min_{\psi} \mathcal{L}(\psi^k, w^k, g^{k+1}, y^k, v^k) \\ w^{k+1} = \arg \min_{\psi} \mathcal{L}(\psi^{k+1}, w^k, g^{k+1}, y^k, v^k) \end{cases} \quad (10)$$

then updates the Lagrange multipliers:

$$\begin{cases} y^{k+1} = y^k + \beta \mu (Fg^{k+1} - M \circ \psi^{k+1}) \\ v^{k+1} = v^k + \beta \eta (w^{k+1} - DF^{-1}\psi^{k+1}) \end{cases} \quad (11)$$

where $\beta \in (0, (\sqrt{5} + 1)/2)$ is an appropriately chosen step length. Fortunately, all the subproblems in Eq. (10) have closed-form solutions (see Appendix A for the derivations), so the iteration can be effectively implemented and outputs the estimated signal $\hat{x} = F^{-1}\hat{\psi}$. Because phase retrieval is a non-convex problem, the ADM algorithm cannot be guaranteed to converge. Nevertheless, ADM often performs better than existing phase retrieval algorithms [43].

We evaluated the effect of TV regularization by comparing the PSNR of the measurements before and after TV regularization. Two hundred measurements were tested, with random sampling ratios in the range of 0–100%. The comparisons are shown in Fig. 2. The quality of the recovered image is significantly improved after TV regularization at the same sampling ratio. In other words, fewer measurements are required to achieve the same quality in the TV regularization method. If the PSNR criterion for imaging quality is 20 dB, only ~20% sampling is needed, but the measurement without TV regularization requires ~80% sampling.

However, uniform random sampling does not consider the energy distribution of the actual image in the Fourier domain. Because the spectral energy of most natural images is centralized to a low frequency, the sampling mask should have more dense sampling near the center to match the energy distribution. A similar sampling scheme has been applied in MRI imaging [44]. Here, a polynomial variable-density sampling is tested, and the sampling obeys the following probability density function:

$$f(r) = \frac{p^2 + 3p + 2}{2\pi} (1 - r)^p \quad (12)$$

where $r \in [0, 1)$ is the normalized distance of the sampling point from the spectrum center. The sampling density decreases with the increase of radius r , and the exponential $p \geq 1$ control the degree of the sampling density near the center. The larger the p , the denser the sample of the central part. The coefficient $p^2 + 3p + 2/2\pi$ make the integral over the entire space is equal to 1. The recovered images are shown in Fig. 2(b), which demonstrates that using variable-density sampling instead of uniform sampling further improves the image quality.

3.2. Total variation regularization for Fourier ptychography

In FP imaging, spectrum sampling is achieved by the relative shifting of the aperture and spectrum. For the diffraction-limited imaging system, the actual sampling pattern is determined by the shape of the aperture, which is generally circular or a regular polygon. Multiple measurements acquired by aperture scanning or multiple cameras should be involved in image reconstruction. Thus, TVFP can be realized by extending the objective function of Eq. (7) to multiple measurements:

$$\min_x \tau \|Dx\|_1 + \frac{1}{2} \sum_{i=1}^N \left\| F^{-1}(P \circ S_i Fx) - b_i \right\|_2^2 \quad (13)$$

where N is the total number of measurements, and S_i is the i th shifting operator, meaning that $S_i \psi$ is equivalent to $\psi(\mathbf{u} - \mathbf{c}_i)$. The augmented Lagrangian function of the problem is given by:

$$\begin{aligned} \mathcal{L}_{TVFP}(\psi, w, g_i, y_i, v) = & \tau \|w\|_1 + \frac{1}{2} \sum_{i=1}^N \|g_i - b_i\|_2^2 \\ & + \frac{\mu}{2} \sum_{i=1}^N \left\| Fg_i - P \circ S_i \psi + \frac{1}{\mu} y_i \right\|_2^2 \\ & + \frac{\eta}{2} \|w - DF^{-1}\psi + \frac{1}{\eta} v\|_2^2 \end{aligned} \quad (14)$$

The ADM iteration scheme is then applied for Eq. (14) (see the Appendix A for the derivations). The TVFP is summarized below as Algorithm 1. In Algorithm 1, the initial value of ψ can be set to the Fourier spectrum of the low-resolution image measured at the center, and the initial value of other variables can be set to zero.

To demonstrate the effect of TV regularization for FP reconstruction, the results recovered by Gauss-Newton method, a conventional FP algorithm, and the proposed TVFP algorithm were compared. In the simulations, a resolution chart with 512×512 pixels was used as the amplitude-only object, and its physical size was specified as 64 mm \times 64 mm. An image system located at 15 m away from the object provided the measurements. The focal length of the image system was 400 mm, and the aperture diameter was 10 mm. The illumination wavelength was

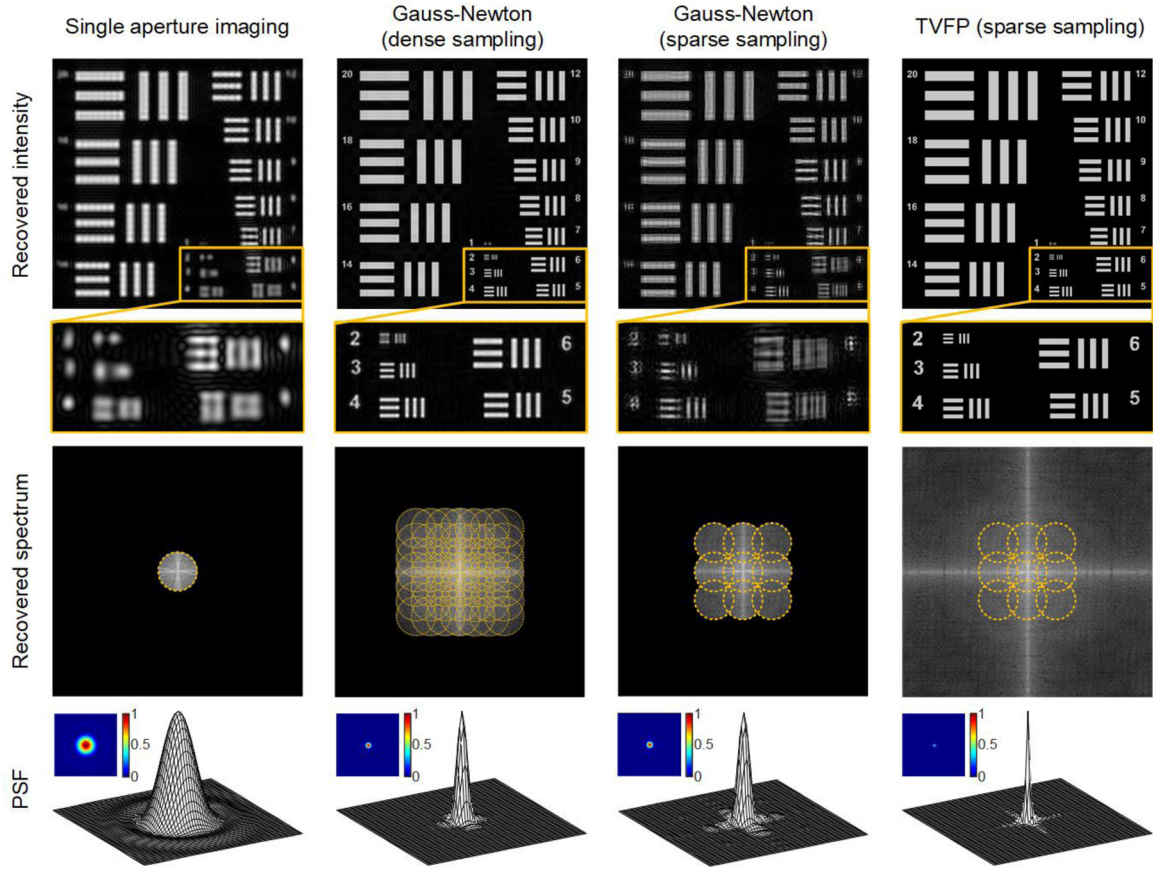


Fig. 3. Reconstructions of the resolution chart with Gauss-Newton and the proposed TVFP method. The single aperture imaging result is shown in first column for reference. The second row show the reconstructed spectrum where the dotted circles indicate the sampled spectral region by aperture. The third row show corresponding PSF. Note that Gauss-Newton can only recover the spectrum in the aperture sampling region. The sparse spectral sampling causes the serious side-lobe effect of PSF. By contrast, TVFP can recover the spectrum outside the sampling region and obtain high-quality image.

Algorithm 1: TVFP algorithm.

Input: Pupil function P , shifting operators S_i , magnitude measurements b_i
Output: Recovered object x

- 1: **Initialization** $\psi^0, w^0 = 0, y_i^0 = 0, v^0 = 0, \lambda, \mu, \eta, \beta$
- 2: **while** stopping conditions are not satisfied **do**
- 3: $g_i^{k+1} = \arg \min_g \mathcal{L}_{TVFP}(\psi^k, w^k, g_i^k, y_i^k, v^k)$
- 4: $\psi^{k+1} = \arg \min_\psi \mathcal{L}_{TVFP}(\psi^k, w^k, g_i^{k+1}, y_i^k, v^k)$
- 5: $w^{k+1} = \arg \min_w \mathcal{L}_{TVFP}(\psi^{k+1}, w^k, g_i^{k+1}, y_i^k, v^k)$
- 6: $y_i^{k+1} = y_i^k + \beta \mu (F g_i^{k+1} - P \circ S_i \psi^{k+1})$
- 7: $v^{k+1} = v^k + \beta \eta (w^{k+1} - D F^{-1} \psi^{k+1})$
- 8: $k \leftarrow k + 1$
- 9: **end while**
- 10: **return** $x = F^{-1} \psi$

assumed to be 532 nm. According to the above parameters, the aperture diameter could be calculated as 80 pixels. The recovered results are shown in Fig. 3.

The first column in Fig. 3 shows the captured image of the center aperture and corresponding spectrum obtained by phase retrieval. Owing to the limited size of the aperture, the point spread function (PSF) has a broad width. Then a 7×7 dense sampling is adopted to expand the passband and narrow the PSF. The spacing between each camera was 3.75 mm, corresponding to a 44% overlapping ratio. The maximum aperture scanning range was 22.5 mm in both the horizontal and vertical directions. The image recovered by Gauss-Newton method becomes clear as shown in the second column of Fig. 3, and the width of the PSF is narrower than that from the single aperture. However, if the number of measurements is reduced to 3×3 , and overlapping ratio is reduced

to 25%, which results a sparse aperture arrangement, the side-lobe of the PSF degrades the image, as shown in the third column of Fig. 3. By contrast, the results recovered by the TVFP algorithm using 3×3 sparse sampling are shown in the last column of Fig. 3. Owing to the introduction of TV regularization, phase consistency was preserved even at low overlapping ratio. The spectrum outside the sampling region could also be recovered, and the side-lobe of the PSF was eliminated. Thus, in the same sampling mode, the resolution of the image recovered by the TVFP algorithm is higher than that of the Gauss-Newton method. In other words, the TVFP algorithm can obtain the same quality image with fewer measurements.

3.3. Pupil function estimation

In a practical optical system, the presence of aberrations may result in phase distortions in the pupil function in Eq. (2). The phase distortions can have a severe effect on the fidelity of the imaging system. The pupil function is usually difficult to obtain, which limits the usage of the TVFP algorithm. In ptychographic imaging, Rodenburg et al. [45] proposed the extended ptychographic iterative engine (ePIE), which could recover both the object and the probe function. Ou et al. [12] transferred this method to FP reconstruction, termed embedded pupil function recovery (EPRY). In EPRY, both the Fourier spectrum and the pupil function are simultaneously recovered by alternately updating the spectrum and pupil function estimates in each iteration. In this work, a joint regularization is proposed to reconstruct the phase distortion of the pupil function and we refer to this enhanced version as extended TVFP (eTVFP).

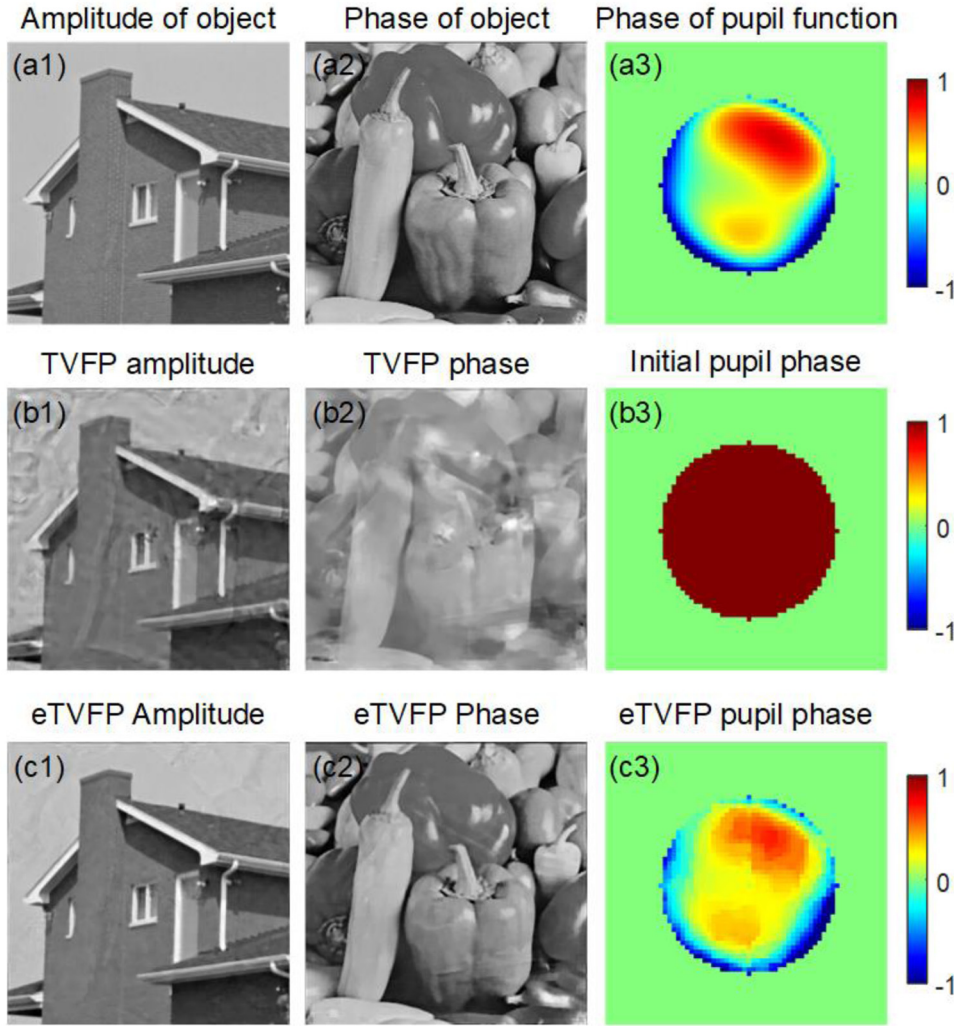


Fig. 4. Reconstructions for TVFP and eTVFP algorithms. (a1) – (a2) The amplitude and phase of the sample. (a3) Phase of the pupil function used to generate simulated measurements. (b1) – (b2) Reconstructed amplitude and phase using the TVFP algorithm which does not correct the aberration. (b3) Initial guess of the pupil phase used in both TVFP and eTVFP algorithm. (c1) – (c2) Reconstructed amplitude and phase using eTVFP algorithm. (c3) Reconstructed pupil function phase, showing a similar distribution as (a3).

The optimization of the pupil function follows the objective function presented in Eq. (13) by changing the optimization variable to the pupil function P , and the TV regularization is also imposed on P to preserve its smoothness. In addition, the initial guess of the pupil function can be introduced into optimization as a support constraint. In the k th iteration, with the knowledge of the reconstructed Fourier spectrum ψ^k and g_i^k from the current iteration, the pupil function can be updated by minimizing the objective:

$$\min_P \tau_p \|DP\|_1 + i_S(P) + \frac{1}{2} \sum_{i=1}^N \|g_i - b_i\|_2^2, \quad (15)$$

s.t. $Fh = P$, $g_i = F^{-1}(S_i \psi \circ Fh)$

where h is the inverse Fourier transform of P and represents the PSF of the imaging system. $i_S(\cdot)$ is the indicator function with $i_S(P) = 0$ if $P \in S$, and $i_S(P) = +\infty$ otherwise. S represents the signal set satisfying the support constraint and modular constraint, that is, $S = \{P : |P(i)| = 0, \text{ if } i \notin D \text{ and } |P| \leq 1\}$. Unfortunately, Eq. (15) does not have a closed-form solution. ADM is applied again to seek the optimal P in each iteration. The update of the pupil function includes the following steps (see Appendix B for the derivations).

Step 1: Update PSF h

$$Fh^{t+1} = \frac{\mu_p \sum_{i=1}^N (S_i \psi)^* \circ F \left(g_i + \frac{1}{\mu_2} y_i \right) + \gamma \left(P + \frac{1}{\gamma} z \right)}{\mu_p \sum_{i=1}^N |S_i \psi|^2 + \gamma I} \quad (16)$$

Step 2: Update the pupil function P

$$FP' = \frac{F \left[\xi D^T \left(m + \frac{1}{\xi} u^t \right) + \gamma \left(Fh^{t+1} - \frac{1}{\gamma} z^t \right) \right]}{\xi (I_h + I_v) + \gamma I} \quad (17)$$

we denote $\mathcal{P}_S(\cdot)$ as the projection onto S . Then update P^{t+1} :

$$P^{t+1} = \mathcal{P}_S[P'(i)] = \begin{cases} \text{sgn}(P'_i), & i \in D \text{ and } |P'_i| > 1 \\ 0, & i \notin D \end{cases} \quad (18)$$

Step 3: Update the auxiliary variables m and Lagrange multipliers u , z :

$$m^{t+1} = \text{sgn} \left(DP^{t+1} - \frac{1}{\xi} u^t \right) \times \max \left(\left| DP^{t+1} - \frac{1}{\xi} u^t \right| - \frac{\tau_p}{\xi}, 0 \right) \quad (19)$$

$$u^{t+1} = u^t + \beta \xi (m^{t+1} - DFh^{t+1}) \quad (20)$$

$$z^{t+1} = z^t + \beta \gamma (P^{t+1} - Fh^{t+1}) \quad (21)$$

The pupil function is minimized by iteratively performing the processes from Eq. (16) to Eq. (21), in which the Fourier spectrum ψ and g_i remain unchanged throughout the sub-iterative process. The experience demonstrates that five sub-iterations are sufficient for convergence. The eTVFP algorithm is summarized in Algorithm 2. The Matlab codes of TVFP and eTVFP are available in GitHub [46].

A numerical simulation of imaging with an aberration system is presented. Two 256×256 -pixel images with $1 \mu\text{m}$ pixel pitch were used as

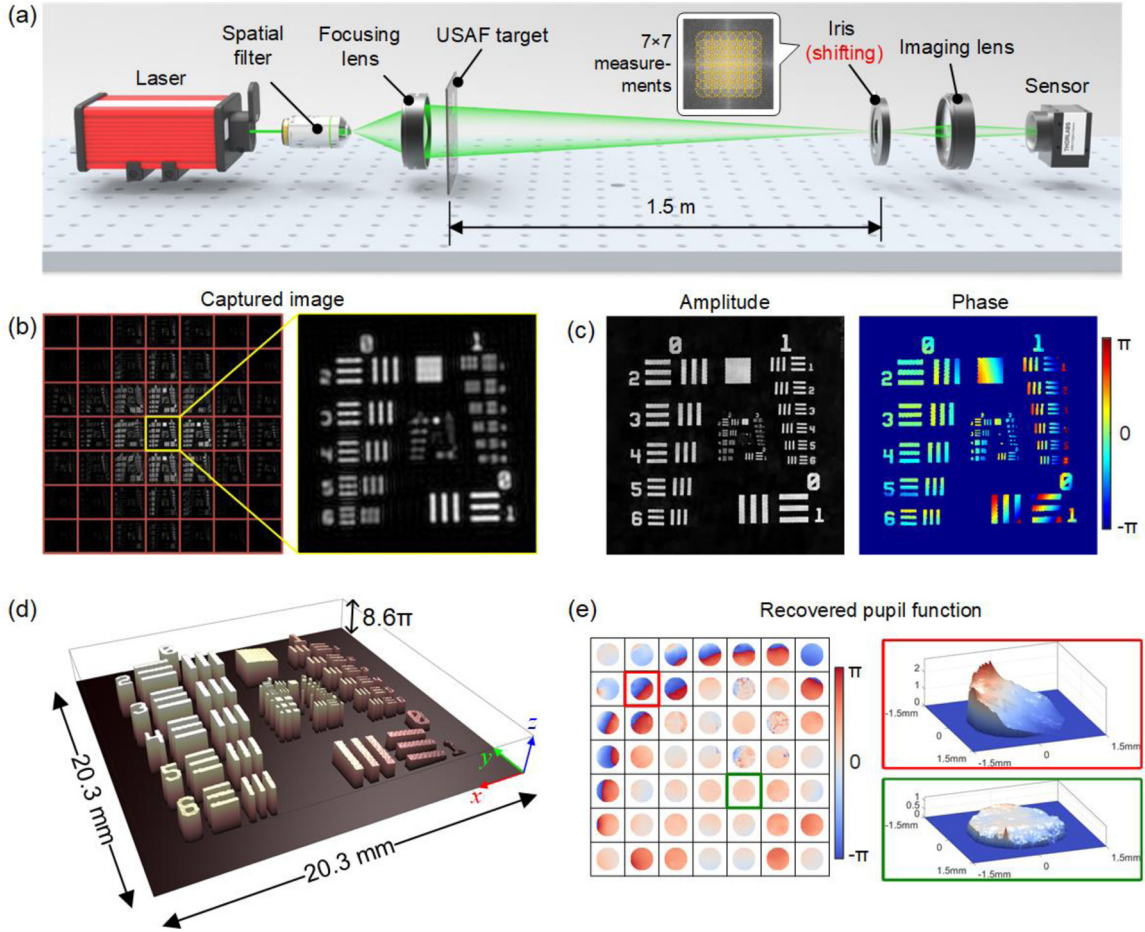


Fig. 5. Experiments for imaging an USAF target by TVFP method. (a) Experimental setup of long-range imaging. (b) The full measurement mosaic acquired by scanning the aperture. (c) The recovered amplitude and phase. (d) Pseudo-3D rendering for the unwrapped phase. (e) The recovered pupil function mosaic and the unwrapped phase of the pupil function close to the edge and center. The diameter of each pupil is 2 mm, and the spacing of adjacent pupils is 1 mm, which means the overlapping ratio is 50 % in this full sampling mode. The pupil function close to the edge shows more severe aberration than close to center.

Algorithm 2: eTVFP algorithm.

Input: Pupil support D , shifting operators S_i , measurement vectors b_i
Output: Recovered object x , pupil function P

```

1: Initialization  $\psi^0, w^0 = 0, y_i^0 = 0, u^0 = 0, z^0 = 0, v^0 = 0, \lambda, \lambda_p, \mu, \mu_p, \eta, \eta_p, \gamma, \beta$ 
2: while not converged do
3:    $g_i^{k+1} = \arg \min_{\psi} \mathcal{L}_{TVFP}(\psi^k, w^k, g_i^k, y_i^k, v^k)$ 
4:    $\psi^{k+1} = \arg \min_{\psi} \mathcal{L}_{TVFP}(\psi^k, w^k, g_i^{k+1}, y_i^k, v^k)$ 
5:   for  $t = 1$  to 5
6:     update  $h^{t+1}$  according to Eq. (16)
7:     update  $P^{t+1}$  according to Eq. (17) and (18)
8:     update  $m^{t+1}$  according to Eq. (19)
9:      $u^{t+1} = u^t + \beta \xi(m^{t+1} - DF^t h^{t+1})$ 
10:     $z^{t+1} = z^t + \beta \gamma (P^{t+1} - F^t h^{t+1})$ 
11:   end
12:    $w^{k+1} = \arg \min_{w} \mathcal{L}_{TVFP}(\psi^{k+1}, w^k, g_i^{k+1}, y_i^k, v^k)$ 
13:    $y_i^{k+1} = y_i^k + \beta \mu (F \psi^{k+1} - P \circ S_i \psi^{k+1})$ 
14:    $v^{k+1} = v^k + \beta \eta (w^{k+1} - DF^{-1} \psi^{k+1})$ 
15:    $k \leftarrow k + 1$ 
16: end while
17: return  $x = F^{-1} \psi, P$ 

```

the amplitude and phase of the observed object, as shown in Fig. 4(a1) – (a2). The imaging system had an f-number $f/12$ corresponding to a circular pupil function with a radius of 20 pixels, and the phase of the pupil function represented the wave aberration, as shown in Fig. 4(a3).

A 5×5 grid-sampled images with 25% overlap were used to reconstruct the high-resolution image. In both algorithms, the initial guess of the pupil function is circular with a constant phase, as shown in Fig. 4(b3), and both algorithms were tested for 500 iterations. Because the TVFP algorithm does not correct the aberration, the recovered amplitude and phase shown in Fig. 4(b1) – (b2) have undesired quality, and the images show crosstalk between the amplitude and phase. In comparison, the eTVFP reconstructions shown in Fig. 4(c1) – (c2) almost eliminate the crosstalk and significantly improve the imaging quality, while the phase of the pupil function could be estimated, as shown in Fig. 4(c3).

4. Experiments

The experimental setup for data acquisition is shown in Fig. 5(a). The spatially filtered laser with a wavelength of 532 nm passed through the focusing lens, forming a convergent spherical wave. A USAF target was placed 1.5 m away from the imaging system and illuminated by the converging spherical wave. The imaging system consisted of a 75 mm focal length lens, an iris with a 30 mm max aperture diameter, and a QHY163 CMOS sensor with 3.8 μm pixel pitch. The iris mounted on a 2D translation stage was used to simulate movement of a large f-number lens. During the image acquisition process, the iris was set to a 2 mm, corresponding to a $f/37.5$ imaging system. A 7×7 grid of images with 1 mm aperture spacing between adjacent images was recorded, resulting in a synthetic aperture with an 8 mm diameter, which corresponds to a

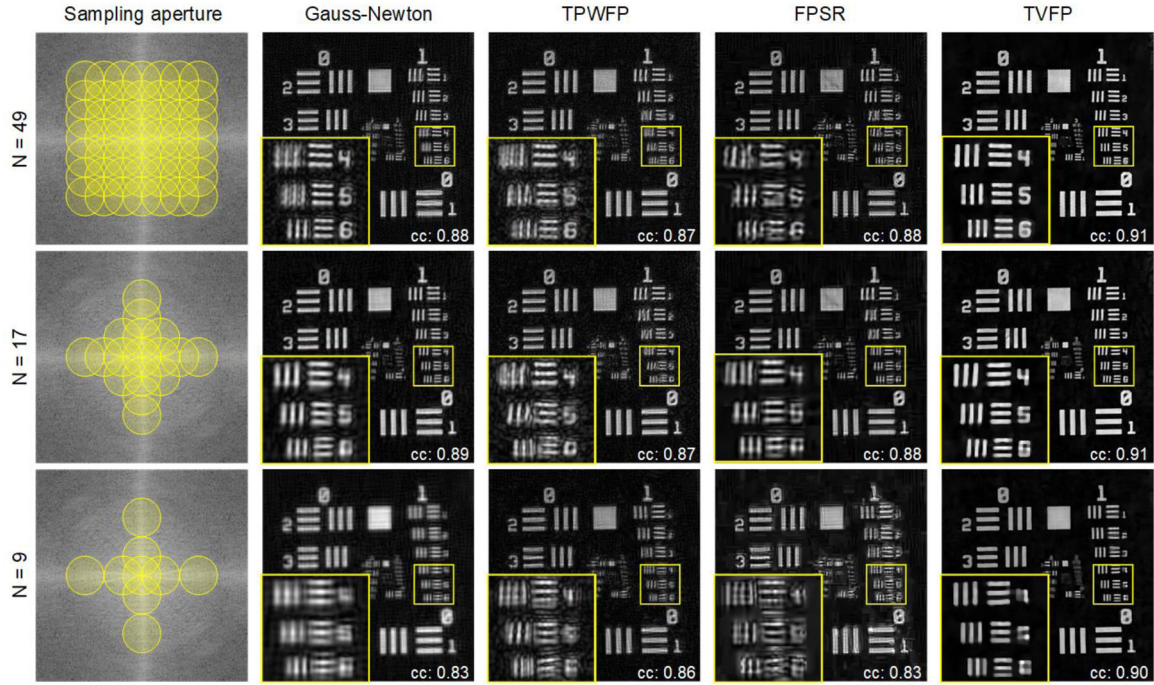


Fig. 6. Reconstructed amplitude for USAF by the four methods (Gauss-Newton, TPWFP, FPSR, and eTVFP) for different numbers of measurements. The first column shows the aperture arrangement with different aperture number. The other columns show the reconstruction results of different methods under the corresponding aperture number. The yellow boxed regions are zoomed in for clarity. The value of correlation coefficient is annotated in the lower right corner of the images.

$f/9.4$ imaging system. A set of images with different exposure times was captured at each aperture position, to merge into a high-dynamic-range image, which are shown in Fig. 5(b). The resolution of the USAF target is limited to 1.78 line pairs per millimeter (lp/mm) (Group 0, Element 6) in the center observed image.

The reconstructed amplitude and phase are shown in Fig. 5(c). By aperture synthesis, the spatial resolution was increased to 5.04 lp/mm (Group 2, Element 3), giving a factor of $2.8 \times$ improvement in resolution. The pseudo-3D results reconstructed by the proposed eTVFP are shown in Fig. 5(d), where the x - and y -axes represent the spatial coordinates of the object, and the z -axis represents the unwrapped phase. Subject to the phase of the illumination light and the tilt of the sensor, the phase shape of the target is bent rather than flat. The phase range reaches 8.6π . Since the wave aberration is different at each position of the imaging lens, the pupil function is estimated for each measured position. As shown in Fig. 5(e), the top left corner has much more severe aberration compared with the other positions. This asymmetry may be caused by the central aperture not aligning with the center of the imaging lens.

The eTVFP was performed for different numbers of measurements to evaluate the potential in the sparse sampling scheme. Three state-of-the-art FP methods, namely Gauss-Newton method [11], truncated Poisson Wirtinger Fourier ptychographic (TPWFP) [15] and Fourier ptychographic with sparse representation (FPSR) [26] were chosen for comparison. In the sparse aperture case, several images were selected from 49 measurements for reconstruction, while the others were discarded. The corresponding aperture arrangement are shown in the leftmost column of Fig. 6, following the variable-density sampling with denser sampling near the center of spectrum. Note that in the case of 9 measurements, which means only 18.4% of the measurements are utilized, the peripheral 4 measurements have no spectral overlap with the centric 5 measurements. Gauss-Newton reconstructions lack prior information of the reconstructed image, which results in blurry reconstruction. The TPWFP and FPSR methods have similar performances because they are based on the Wirtinger flow method and maximum-likelihood estimation for the noisy observation model. Although FPSR introduces the Block Matching 3-D (BM3D) method to implement the sparse representation of the

reconstructed image, the sparse representation is not incorporated into the optimization model, which may lead to iteration cannot converge.

To quantitatively evaluate the quality of reconstructed images, a full aperture image, corresponding to a $f/2.95$ imaging system, is captured as a reference image. Due to the inconsistency of the reconstructed image in pixel value level, Pearson correlation coefficient is adopted as a metric instead of PSNR. Pearson correlation coefficient can keep the same for any linear amplification and bias of image pixel value. The value of correlation coefficient is annotated in the lower right corner of the images in Fig. 6. By contrast, eTVFP provides more image details and higher correlation coefficient than other methods.

5. Conclusion

In this work, a TV regularization-based Fourier ptychographic reconstruction, called the TVFP method, is proposed for long-range imaging. To reduce redundant measurements, TV regularization is introduced into the framework of FP reconstruction. The prior knowledge imposed by TV regularization mitigates the ill-posedness of phase retrieval and potentially allows recovery of the complete spectrum. With a joint regularization, the phase distortion at the pupil function could also be recovered in our proposed eTVFP method. The analysis for the sampling ratio and overlapping ratio demonstrate the outstanding performance of TVFP in sparse sampling.

Since TV regularization bases on an assumption that the gradient distribution of the underlying image is sparse, the TVFP method is more suitable for imaging the target which is piecewise smooth in appearance. Failures of the TV regularization might be suspected if the target has intricate textures or jumpy phase such that the details of the recovered image are smoothed out. The problem can be alleviated by appropriately reducing regularization weight and increasing the overlapping ratio. On the contrary, if we know by prior knowledge that the target has a simple structure, the algorithm has faster convergence speed by increasing regularization weight while the overlapping ratio could be reduced. In certain case, the proposed TVFP allows for a zero overlapping ratio, showing the potential for single-shot FP imaging by using a camera ar-

ray instead of inherent aperture scanning in traditional FP imaging, so that the observation of rapidly changing scenes can be realized.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Jiachen Wu: Writing – original draft, Methodology, Visualization.
Feng Yang: Conceptualization, Validation, Data curation. **Liangcai Cao:** Supervision, Funding acquisition, Writing – review & editing.

Funding

This work was supported by National Nature Science Foundation of China (61775117, 61827825) and the Tsinghua University Initiative Scientific Research Program (20193080075).

Acknowledgment

The authors thank Miss Yaoyao Li for collecting data during the COVID-19.

Appendix A

Derivation of subproblems in TVFP

The closed-form solutions of the subproblems (10) are given in the following derivation.

1. Solution of g -subproblem.

The subproblem to minimize \mathcal{L} over g is simplified by omitting the irrelative terms:

$$\min_g \frac{1}{2} \|g - b\|_2^2 + \frac{\mu}{2} \left\| Fg - M \circ \psi + \frac{1}{\mu} y \right\|_2^2 \quad (\text{A.1})$$

Due to the exist of taking modulus operator, $\|g - b\|_2^2$ is nonconvex. The least square method is not available. Let $s = F^{-1}(M \circ \psi - y/\mu)$, then the augmented Lagrangian involving g is expressed as

$$\begin{aligned} & \frac{1}{2} \|g - b\|_2^2 + \frac{\mu}{2} \|g - s\|_2^2 \\ & \geq \frac{1}{2} \|g - b\|_2^2 + \frac{\mu}{2} \|g - |s|\|_2^2 \\ & = \frac{1 + \mu}{2} \left(|g| - \frac{b + \mu|s|}{1 + \mu} \right)^2 + C \end{aligned} \quad (\text{A.2})$$

where the '=' in the inequality holds if $\text{sgn}(g) = \text{sgn}(s)$, and C is a constant independent of g . As a result, the augmented Lagrangian obtains the minimum when g satisfied

$$g = \frac{b + \mu|s|}{1 + \mu} \text{sgn}(s) \quad (\text{A.3})$$

2. Solution of ψ -subproblem.

The subproblem to minimize \mathcal{L} over ψ is simplified by omitting the irrelative terms:

$$\min_{\psi} \frac{\mu}{2} \left\| Fg - M \circ \psi + \frac{1}{\mu} y \right\|_2^2 + \frac{\eta}{2} \left\| w - DF^{-1}\psi + \frac{1}{\eta} v \right\|_2^2 \quad (\text{A.4})$$

The minimization of Eq. (A.4) is a least squares problem. The solution to the ψ subproblem could be obtained by setting the partial derivative respect to ψ equal to zero:

$$\psi = \frac{\mu M^* \circ \left(Fg + \frac{1}{\mu} y \right) + \eta F D^T \left(w + \frac{1}{\eta} v \right)}{\mu |M|^2 + \eta |H|^2} \quad (\text{A.5})$$

where H is the transfer function of difference operator, that is $D = F^{-1} H F$.

3. Solution of w -subproblem.

The minimization of \mathcal{L} with respect to w is a soft threshold problem:

$$\min_w \tau \|w\|_1 + \frac{\eta}{2} \left\| w - DF^{-1}\psi + \frac{1}{\eta} v \right\|_2^2 \quad (\text{A.6})$$

which has a closed-form solution:

$$w = \text{sgn} \left(DF^{-1}\psi - \frac{1}{\eta} v \right) \cdot \max \left(\left| DF^{-1}\psi - \frac{1}{\eta} v \right| - \frac{\tau}{\eta}, 0 \right) \quad (\text{A.7})$$

The process of ADM algorithm is updating the above variables in sequence according to Eq.(A.3), (A.5) and (A.7) in each iteration. Note that the update order could be arbitrary, so which variables need to be initialized are adjusted according to the first updated variable.

The Fourier ptychography extends the phase retrieval from incomplete spectrum sampling to multiple measurements. The sampling mask M is replaced by the aperture function P , and a shifting operation S_i is introduced into objective function. The augmented Lagrangian function is

$$\begin{aligned} \mathcal{L}_{TVFP}(\psi, w, g_i, y_i, v) = & \tau \|w\|_1 + \frac{1}{2} \sum_{i=1}^N \|g_i - b_i\|_2^2 \\ & + \frac{\mu}{2} \sum_{i=1}^N \left\| Fg_i - P \circ S_i \psi + \frac{1}{\mu} y_i \right\|_2^2 \\ & + \frac{\eta}{2} \left\| w - DF^{-1}\psi + \frac{1}{\eta} v \right\|_2^2 \end{aligned} \quad (\text{A.8})$$

Similar to the derivation of phase retrieval from incomplete spectrum sampling, the minimization of Eq. (A.8) is updating following variables in turns.

$$g_i = \frac{b_i + \mu|S_i|}{1 + \mu} \text{sgn}(S_i), \quad S_i = F^{-1}(P \circ S_i \psi) - y_i/\mu \quad (\text{A.9})$$

$$\psi = \frac{\mu \sum_{i=1}^N \left\{ S_i^T \left[P^* \circ \left(Fg_i + \frac{1}{\mu} y_i \right) \right] \right\} + \eta F D^T \left(w + \frac{1}{\eta} v \right)}{\mu \sum_{i=1}^N (S_i^T |P|^2) + \eta |H|^2} \quad (\text{A.10})$$

$$w = \text{sgn} \left(DF^{-1}\psi - \frac{1}{\eta} v \right) \cdot \max \left(\left| DF^{-1}\psi - \frac{1}{\eta} v \right| - \frac{\tau}{\eta}, 0 \right). \quad (\text{A.11})$$

Appendix B

Derivation of subproblems in eTVFP

The objective function of pupil function estimation is

$$\begin{aligned} \min \tau_p \|DP\|_1 + i_S(P) + \frac{1}{2} \sum_{i=1}^N \|g_i - b_i\|_2^2, \\ \text{s.t. } Fh = P, g_i = F^{-1}(S_i \psi \circ Fh). \end{aligned} \quad (\text{B.1})$$

The augmented Lagrangian function for problem (B.1) is

$$\begin{aligned} \mathcal{L}_p(P, h, m, g_i, y_i, u, z) \\ = & \tau_p \|m\|_1 + \frac{1}{2} \sum_{i=1}^N \|g_i - b_i\|_2^2 + i_S(P) \\ & + \frac{\mu_p}{2} \sum_{i=1}^N \|g_i - F^{-1}(S_i \psi \circ Fh) + \frac{1}{\mu_p} y_i\|_2^2 \\ & + \frac{\xi}{2} \|m - DP + \frac{1}{\xi} u\|_2^2 \\ & + \frac{\gamma}{2} \|P - Fh + \frac{1}{\gamma} z\|_2^2 \end{aligned} \quad (\text{B.2})$$

Since Eq. (B.2) shares the variables g_i and y_i with Eq. (A.8), the minimization of Eq. (B.2) could be realized by updating only P, h, m, u and z .

1. Solution of h -subproblem.

The subproblem to minimize \mathcal{L}_p over h is simplified by omitting the irrelative terms:

$$\min_h \frac{\mu_p}{2} \sum_{i=1}^N \|g_i - F^{-1}(S_i \psi \circ Fh) + \frac{1}{\mu_p} y_i\|_2^2 + \frac{\gamma}{2} \|P - Fh + \frac{1}{\gamma} z\|_2^2 \quad (\text{B.3})$$

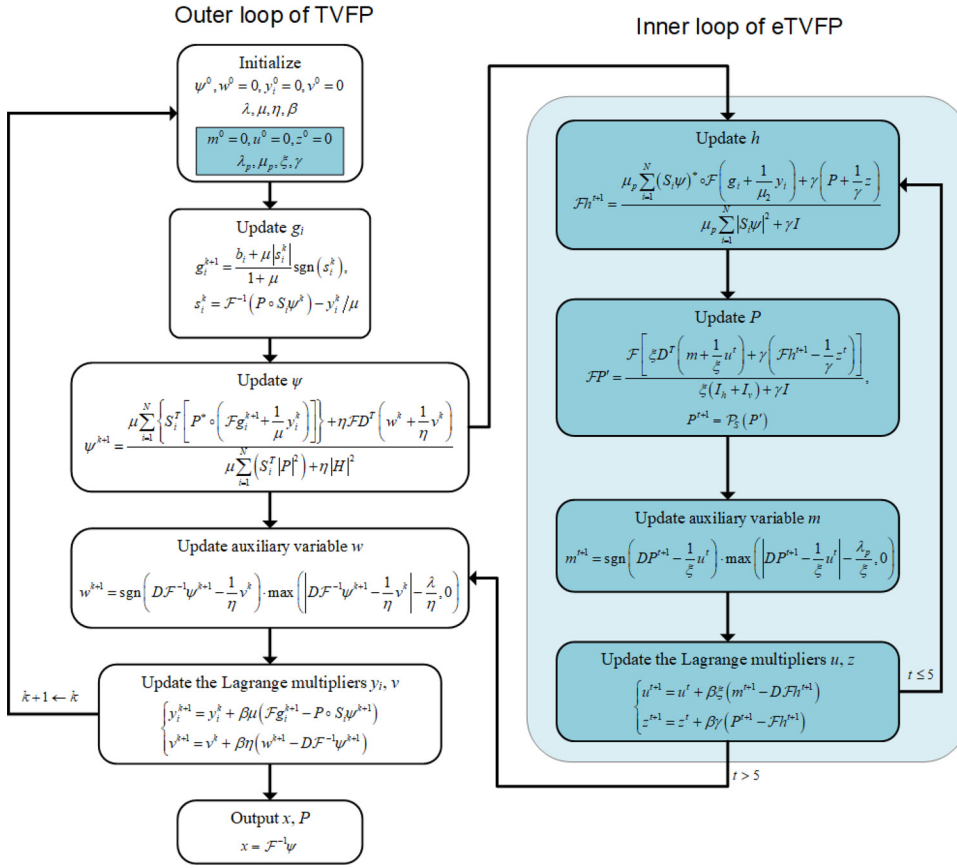


Fig. B. 1. Flow chart of TVFP and eTVFP algorithm.

which is a least squares problem. The solution to the h subproblem could be obtained by setting the partial derivative respect to ψ equal to zero:

$$Fh = \frac{\mu_p \sum_{i=1}^N (S_i \psi)^* \circ F \left(g_i + \frac{1}{\mu_p} y_i \right) + \gamma \left(P + \frac{1}{\gamma} z \right)}{\mu_p \sum_{i=1}^N |S_i \psi|^2 + \gamma I} \quad (\text{B.4})$$

where I is an identity matrix. The reason for keeping Fourier operator in the left side of Eq. (B.4) is that the Fh will be used as a combined term in the updating of P .

2. Solution of P -subproblem.

The subproblem to minimize \mathcal{L}_p over P is as following:

$$\min_P \frac{\xi}{2} \left\| m - DP + \frac{1}{\xi} u \right\|_2^2 + \frac{\gamma}{2} \left\| P - Fh + \frac{1}{\gamma} z \right\|_2^2 \quad (\text{B.5})$$

The solution of P -subproblem is given by

$$FP' = \frac{F \left[\xi D^T \left(m + \frac{1}{\xi} u \right) + \gamma \left(Fh - \frac{1}{\gamma} z \right) \right]}{\xi |H|^2 + \gamma I} \quad (\text{B.6})$$

Then take the inverse Fourier transform of right side of Eq. (B.6) to obtain the updated P' . In addition, the size and the shape of aperture could be introduced to pupil estimation as a support constraint. Denote $\mathcal{P}_S(\cdot)$ the projection (in Euclidean norm) onto S . To satisfy the support constraints, the updated P should be further projected onto S , that is $P = \mathcal{P}_S(P')$.

3. Solution of m -subproblem.

The minimization of \mathcal{L}_p with respect to m is a soft threshold problem:

$$\min_m \tau_p \|m\|_1 + \frac{\xi}{2} \left\| m - DP + \frac{1}{\xi} u \right\|_2^2 \quad (\text{B.7})$$

which has a closed-form solution:

$$m = \text{sgn} \left(DP - \frac{1}{\xi} u \right) \cdot \max \left(\left| DP - \frac{1}{\xi} u \right| - \frac{\tau_p}{\xi}, 0 \right) \quad (\text{B.8})$$

The flow chart of TVFP and eTVFP algorithm is shown in Fig. B. 1. For TVFP algorithm, only the steps on the white background are performed. If the pupil function is needed to be estimated, the cyan part corresponding to the inner loop of eTVFP algorithm is also needed to be performed.

References

- [1] Ryle M, Hewish A. The synthesis of large radio telescopes. *Mon Not R Astron Soc* 1960;120:220–30.
- [2] Bashkansky M, Lucke RL, Funk E, Rickard L, Reintjes J. Two-dimensional synthetic aperture imaging in the optical domain. *Opt Lett* 2002;27:1983–5.
- [3] Miller NJ, Dierking MP, Duncan BD. Optical sparse aperture imaging. *Appl Opt* 2007;46:5933–43. doi:10.1364/AO.46.005933.
- [4] Brown WM. Synthetic aperture radar. *IEEE Trans Aerosp Electron Syst* 1967;217–29.
- [5] Bouman KL, Johnson MD, Zoran D, Fish VL, Doeleman SS, Freeman WT. Computational imaging for VLBI image reconstruction. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; 2016. p. 913–922.
- [6] Zheng G, Horstmeyer R, Yang C. Wide-field, high-resolution Fourier Ptychographic microscopy. *Nat Photonics* 2013;7:739.
- [7] Ou X, Horstmeyer R, Yang C, Zheng G. Quantitative phase imaging via Fourier Ptychographic microscopy. *Opt Lett* 2013;38:4845–8.
- [8] Zuo C, Sun J, Li J, Asundi A, Chen Q. Wide-field high-resolution 3d microscopy with fourier Ptychographic diffraction tomography. *Opt Lasers Eng* 2020;128:106003.
- [9] Dong S, Horstmeyer R, Shiradkar R, Guo K, Ou X, Bian Z, et al. Aperture-scanning Fourier Ptychography for 3D refocusing and super-resolution macroscopic imaging. *Opt Express* 2014;22:13586–99.
- [10] Holloway J, Asif MS, Sharma MK, Matsuda N, Horstmeyer R, Cossairt O, et al. Toward long distance sub-diffraction imaging using coherent camera arrays. *IEEE Trans Comput Imaging* 2016;2:251–65. doi:10.1109/TCI.2016.2557067.
- [11] Holloway J, Wu Y, Sharma MK, Cossairt O, Veeraraghavan A. SAVI: Synthetic apertures for long-range, subdiffraction-limited visible imaging using Fourier Ptychography. *Sci Adv* 2017;3:e1602564. doi:10.1126/sciadv.1602564.
- [12] Ou X, Zheng G, Yang C. Embedded pupil function recovery for Fourier Ptychographic microscopy. *Opt Express* 2014;22:4960–72.
- [13] Tian L, Li X, Ramchandran K, Waller L. Multiplexed coded illumination for Fourier Ptychography with an LED array microscope. *Biomed Opt Express* 2014;5:2376–89.
- [14] Bian L, Suo J, Zheng G, Guo K, Chen F, Dai Q. Fourier Ptychographic reconstruction using Wirtinger flow optimization. *Opt Express* 2015;23:4856–66.

- [15] Bian L, Suo J, Chung J, Ou X, Yang C, Chen F, et al. Fourier ptychographic reconstruction using Poisson maximum likelihood and truncated Wirtinger gradient. *Sci Rep* 2016;6:27384.
- [16] Horstmeyer R, Chen RY, Ou X, Ames B, Tropp JA, Yang C. Solving ptychography with a convex relaxation. *New J. Phys.* 2015;17:053044.
- [17] Jagatap G, Chen Z, Nayer S, Hegde C, Vaswani N. Sample efficient fourier ptychography for structured data. *IEEE Trans Comput Imaging* 2019;6:344–57.
- [18] Bolcar MR, Fienup JR. Method of phase diversity in multi-aperture systems utilizing individual sub-aperture control. In: *Unconventional Imaging*; 2005. p. 58960G.
- [19] Guizar-Sicairos M, Fienup JR. Understanding the twin-image problem in phase retrieval. *J Opt Soc Am A* 2012;29:2367–75. doi:10.1364/JOSAA.29.002367.
- [20] Bunk O, Dierolf M, Kynde S, Johnson I, Marti O, Pfeiffer F. Influence of the overlap parameter on the convergence of the ptychographical iterative engine. *Ultramicroscopy* 2008;108:481–7.
- [21] Dong S, Bian Z, Shiradkar R, Zheng G. Sparsely sampled Fourier ptychography. *Opt Express* 2014;22:5455–64.
- [22] Zhao F, Noll DC, Nielsen J-F, Fessler JA. Separate magnitude and phase regularization via compressed sensing. *IEEE Trans Med Imaging* 2012;31:1713–23.
- [23] Yang Z, Zhang C, Xie L. Robust compressive phase retrieval via L1 minimization with application to image reconstruction. *arXiv:1302.0081* 2013.
- [24] Chang H, Lou Y, Ng MK, Zeng T. Phase retrieval from incomplete magnitude information via total variation regularization. *SIAM J Sci Comput* 2016;38:A3672–95.
- [25] Zhang Y, Cui Z, Zhang J, Song P, Dai Q. Group-based sparse representation for Fourier ptychography microscopy. *Opt Commun* 2017;404:55–61.
- [26] Zhang Y, Song P, Zhang J, Dai Q. Fourier ptychographic microscopy with sparse representation. *Sci Rep* 2017;7:8664. doi:10.1038/s41598-017-09090-8.
- [27] Li Z, Wen D, Song Z, Liu G, Zhang W, Wei X. Sub-diffraction visible imaging using macroscopic fourier ptychography and regularization by denoising. *Sensors* 2018;18:3154.
- [28] Sun Y, Xu S, Li Y, Tian L, Wohlberg B, Kamilov US. Regularized fourier ptychography using an online plug-and-play algorithm. In: *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*; 2019. p. 7665–7669.
- [29] Jagatap G, Chen Z, Hegde C, Vaswani N. Sub-diffraction imaging using Fourier ptychography and structured sparsity. In: *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*; 2018. p. 6493–6497.
- [30] Kappeler A, Ghosh S, Holloway J, Cossairt O, Katsaggelos A. PtychNet: CNN based Fourier ptychography. In: *2017 IEEE International Conference on Image Processing (ICIP)*; 2017. p. 1712–1716.
- [31] Nguyen T, Xue Y, Li Y, Tian L, Nehmetallah G. Deep learning approach for Fourier ptychography microscopy. *Opt Express* 2018;26:26470–84.
- [32] Shamshad F, Abbas F, Ahmed A. Deep ptych: Subsampled Fourier ptychography using generative priors. In: *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*; 2019. p. 7720–7724.
- [33] Zhang J, Xu T, Shen Z, Qiao Y, Zhang Y. Fourier ptychographic microscopy reconstruction with multiscale deep residual network. *Opt Express* 2019;27:8612–25. doi:10.1364/OE.27.008612.
- [34] Rudin LI, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms. *Physica D* 1992;60:259–68.
- [35] Ren D, Bostan E, Yeh LH, Waller L. Total-variation regularized Fourier ptychographic microscopy with multiplexed coded illumination. In: *Mathematics in Imaging*; 2017. p. MM3C. 5.
- [36] Tian X. Fourier ptychographic reconstruction using mixed Gaussian–Poisson likelihood with total variation regularisation. *Electron Lett* 2019;55:1041–3.
- [37] Goodman JW. *Introduction to Fourier optics*. 4th ed. New York: W.H. Freeman; 2017.
- [38] Zhou A, Chen N, Wang H, Situ G. Analysis of Fourier ptychographic microscopy with half of the captured images. *J Opt* 2018;20:095701.
- [39] Lustig M, Donoho D, Pauly JM. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magn Reson Med* 2007;58:1182–95.
- [40] Yang J, Zhang Y, Yin W. A fast alternating direction method for TVL1-L2 signal reconstruction from partial Fourier data. *IEEE J Sel Top Signal Process* 2010;4:288–97.
- [41] Sakhaee E, Entezari A. Sparse partial derivatives and reconstruction from partial Fourier data. In: *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*; 2015. p. 3621–3625.
- [42] Boyd S, Parikh N, Chu E, Peleato B, Eckstein J. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found Trends Mach Learn* 2010;3:1–122. doi:10.1561/22000000016.
- [43] Wen Z, Yang C, Liu X, Marchesini S. Alternating direction methods for classical and ptychographic phase retrieval. *Inverse Probl* 2012;28:115010.
- [44] Lustig M, Donoho DL, Santos JM, Pauly JM. Compressed sensing MRI. *IEEE Signal Process Mag* 2008;25:72–82.
- [45] Maiden AM, Rodenburg JM. An improved ptychographical phase retrieval algorithm for diffractive imaging. *Ultramicroscopy* 2009;109:1256–62.
- [46] GitHub: The Matlab codes of TVFP and eTVFP, <https://github.com/THUHOloLab/TVFP>; 2020