



Wavelet-domain autofocusing algorithm for lensless ptychographic imaging[☆]

Li Liu^a, Bailin Zhuang^{a, }, Jinxiang Du^a, Lei Zhong^a, Haoyang Liang^a, Hanqi Chen^a,
Qihang Zhang^a, Yunhao Liu^a, Honggang Gu^{a,b,c,*}, Shiyuan Liu^{a,b,*} 

^a School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, China

^b Optics Valley Laboratory, Wuhan 430074, China

^c Guangdong HUST Industrial Technology Research Institute, Guangdong Provincial Key Laboratory of Manufacturing Equipment Digitization, Dongguan, 523003, China

ARTICLE INFO

Keywords:

Autofocusing algorithm
Wavelet transform
Ptychography
Calibration accuracy
Calibration uncertainty

ABSTRACT

Misalignment of the specimen-detector distance introduces perturbations in lensless ptychographic scanning coordinates, leading to reconstruction artifacts caused by discrepancies between the inversion algorithm and experimental uncertainties. Here, we propose an accurate and efficient autofocusing algorithm to calibrate the specimen-detector axial distance in lensless ptychography imaging online. The proposed algorithm leverages wavelet-domain multidimensional coefficients to quantify reconstruction clarities in virtual planes within the depth of field, which requires neither careful window searching nor parameter fine-tuning during the iteration process and can effectively escape from local optimums to avoid violent convergence oscillation and crosstalk. Simulations and experiments conducted on both amplitude and biological specimens demonstrate that the proposed algorithm achieves robust convergence in nearly hundred iterations, significantly removes reconstruction artifacts, and delivers a several-fold to orders of magnitude improvement in convergence speed, calibration accuracy and convergence uncertainty compared to conventional sharpness-based autofocusing. These advancements substantially broaden the potential applications, including other coherent diffractive imaging such as coded ptychography (coherent modulated imaging), in-line holography, and multi-plane phase retrieval.

1. Introduction

High-resolution imaging often demands precise focusing to capture intricate structures across different planes. Conventional methods typically depend on delicate mechanical systems or modulation templates to maintain focus or extend the depth of field, posing challenges in terms of time overhead, system complexity, as well as generalizability [1–5]. In contrast, lensless imaging eliminates the need for lenses and leverages iterative algorithm-based reconstruction to facilitate real-time refocusing during the imaging process, ensuring high imaging fidelity. As a result, autofocus techniques have become critical for high-resolution reconstruction and calibration of system parameters for lensless imaging [6–10].

More recently, spanning spatial domain [11–13] to frequency domain [14,15], lateral gradient [16,17] to axial gradient difference [18,19], edge extraction [20–22] to information statistics [23,24], considerable autofocusing strategies have been proposed for diverse

types of specimens in lensless in-line holography. These methods typically aim to identify a single peak (unimodality) to evaluate reconstruction clarity and minimize the single-peak range, ensuring sufficient sensitivity and accuracy in refocusing. However, the presence of conjugate twin-image artifacts in single-shot phase retrieval makes achieving a single-peak modality particularly challenging, especially when there are large deviations in the specimen-detector distance. This difficulty, combined with specimen variability, often results in focusing failures. To address these issues, Ren et al. [25] and Wu et al. [26] employed convolutional neural networks to establish a mapping between defocused holograms and focused reconstructed images, which effectively eliminated conjugate twin-image artifacts and enabled rapid autofocusing without requiring the reconstruction of an image stack, even with unknown or approximate initial distances. Despite these advances, deep learning-based methods [25–28] generally demand large training datasets and exhibit limited generalizability across different specimens.

[☆] This article is part of a special issue entitled: 'Medical Measurement' published in Measurement.

* Corresponding authors at: School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, China.

E-mail addresses: hongganggu@hust.edu.cn (H. Gu), shyliu@hust.edu.cn (S. Liu).

Diversity measurement phase retrieval techniques have addressed the major challenge of conjugate twin-image artifacts in the single-shot coherent diffractive imaging, enabling autofocus-based system parameter accurate calibration. Loetgering et al. [29] and Ma et al. [30] introduced sharpness evaluation functions based on total variation and edge detection operators to identify the optimal focal plane through angular spectrum propagation or axial traversals, respectively. Meanwhile, Dou et al. [31] and Cao et al. [32] proposed a statistical gradient optimization method (Tamura Coefficient) for calibrating the specimen-detector distance in ptychographic imaging. Unlike sharpness-based autofocus strategies, this statistical approach leverages the entire spatial field of view without requiring window selection. However, both sharpness-based and statistics-based autofocus methods rely on evaluation functions derived from image grey levels (contrast), making them highly susceptible to noise and prone to local optima or convergence oscillations. Therefore, Ruan et al. [33] proposed a L_p -norm total variation autofocus algorithm, which adaptively adjusts each pixel to enhance the signal-to-noise ratio in low grey-level regions, enabling more accurate specimen reconstructions and parameter calibration. Recently, Li et al. [34] developed an amplitude difference autofocus method using fractional Fourier transform for multi-plane diversity measurement. This approach integrates spatial and Fourier domain information, offering greater noise robustness and improved parameter calibration accuracy compared to methods based solely on spatial information. Despite these advancements in axial distance refinement, limitations remain. Specifically, the information extraction dimensions in these methods are often too homogeneous (typically single or dual), and the weighting parameters between domains are not adaptive, making them highly sensitive to specimen type.

To reduce sensitivity to specimen variability and enhance noise robustness for parameter calibration, an accurate and fast autofocus algorithm based on the wavelet domain is proposed for lensless ptychographic specimen-detector distance calibration. The algorithm leverages wavelet-domain multidimensional information, including approximate, horizontal, vertical, and diagonal coefficients, to quantify reconstruction clarities in virtual planes within the depth of field. Unlike conventional autofocus methods, the proposed algorithm does not require precise window searching or parameter fine-tuning during iterations. Moreover, it integrates spatial information (approximate coefficients) with multi-frequency domain details (horizontal, vertical, and diagonal coefficients), effectively escaping local optima caused by noise and mitigating convergence oscillations and crosstalk. Simulations and experiments on both amplitude and biological specimens, compared with conventional sharpness-based autofocus algorithms [29,33], demonstrate that the proposed algorithm achieves robust convergence in nearly hundred iterations and significantly eliminates reconstruction artifacts. It delivers improvements in convergence speed, calibration accuracy and convergence uncertainty ranging from several-fold to an order of magnitude. These advancements greatly expand its potential applications to other coherent diffractive imaging techniques, including coded ptychography (coherent modulated imaging), in-line holography, and multi-plane phase retrieval.

2. Methodology

2.1. Ptychographic imaging algorithm

In real space, the focused probe \mathbf{P} is driven by the displacement stage to make an overlapping grid (or other trajectories) scan on the specimen surface \mathbf{O} . At the j -th scanning position, the exit-wave $\psi_j(\mathbf{r})$ on the

specimen downstream can be expressed as

$$\psi_j(\mathbf{r}) = \mathbf{P}(\mathbf{r} - \mathbf{R}_j) \times \mathbf{O}(\mathbf{r}). \quad (1)$$

Through the forward model, the exit-wave $\psi_j(\mathbf{r})$ propagates into the reciprocal space captured by the detector. Therefore, the measurement diffraction field $I_j(\mathbf{u})$ relative to the j -th scanning position can be expressed as

$$I_j(\mathbf{u}) = \|\mathbf{F}\psi_j(\mathbf{r})\|^2 + b, \quad (2)$$

where, \mathbf{u} and \mathbf{r} are coordinates in the reciprocal space and real space, respectively. \mathbf{R}_j is the scanning position coordinate at the j -th scanning position. \mathbf{F} is the forward model, including but not limited to angular spectral propagation, Fresnel diffraction, and Fraunhofer diffraction, and b is a positive offset matrix of the mixed noises. In fact, ptychographic phase retrieval is in pursuit of minimization of the Euclidean distance ϵ at all scanning positions

$$\epsilon = \operatorname{argmin}_{\mathbf{P}(\mathbf{r}-\mathbf{R}_j), \mathbf{O}(\mathbf{r})} \sum_j \left\| \sqrt{I_j(\mathbf{u})} - |\mathbf{F}[\mathbf{P}(\mathbf{r}-\mathbf{R}_j)\mathbf{O}(\mathbf{r})]| \right\|^2. \quad (3)$$

Therefore, in the optimization of the reciprocal space, search for the optimal wave-field $\Psi_j^{\text{opt}}(\mathbf{u})$ that minimizes the distance between the measured diffraction field $I_j(\mathbf{u})$ and the wave-field $\Psi_j(\mathbf{u})$ diffracted by exit-wave in the forward model of the Eq (2),

$$\begin{aligned} \Psi_j^{\text{opt}}(\mathbf{u}) &= \mathbf{G}(\Psi_j(\mathbf{u})) = \operatorname{argmin} \left(I_j(\mathbf{u}), |\Psi_j(\mathbf{u})|^2 \right), \\ \operatorname{argmin} \left(I_j(\mathbf{u}), |\Psi_j(\mathbf{u})|^2 \right) &= \frac{1}{2} \left\| \sqrt{I_j(\mathbf{u})} - \Psi_j(\mathbf{u}) \right\|^2. \end{aligned} \quad (4)$$

The Wirtinger gradient [35] of the $\Psi_j(\mathbf{u})$ and the gradient descent operation can be expressed as

$$\begin{aligned} \nabla \mathbf{G}(\Psi_j(\mathbf{u})) &= \frac{1}{2} \operatorname{diag}(\Psi_j(\mathbf{u})) \times \left(1 - \frac{\sqrt{I_j(\mathbf{u})}}{\Psi_j(\mathbf{u})} \right), \\ \Psi_j'(\mathbf{u}) &= \Psi_j(\mathbf{u}) - \alpha_\Psi \nabla \mathbf{G}(\Psi_j(\mathbf{u})), \end{aligned} \quad (5)$$

Note that for a particular update direction, the fidelity term can be regarded as a convex function with respect to α_Ψ , and thus the optimal step size can be computed by finding the extremes. After determining the optimal step size, the optimal wave-field $\Psi_j^{\text{opt}}(\mathbf{u})$ at the j -th scanning position can be expressed as

$$\Psi_j^{\text{opt}}(\mathbf{u}) = \Psi_j(\mathbf{u}) \times \frac{\sqrt{I_j(\mathbf{u})}}{\Psi_j(\mathbf{u})}, \quad (6)$$

and it represents the modulus constraint projection in the inverse reconstruction. After N scanning positions are traversed, the overlapping constraint projections are also alternately completed. This is the Step 1 in Fig. 1.

In the optimization of the real space, when the optimal wave-field $\Psi_j^{\text{opt}}(\mathbf{u})$ is updated, an appropriate error metric function is applied to retrieve the specimen surface optimal exit-wave $\Psi_j^{\text{opt}}(\mathbf{r})$ as

$$\begin{aligned} E &= \left\| \Psi_j^{\text{opt}}(\mathbf{r}) - \mathbf{P}(\mathbf{r} - \mathbf{R}_j)\mathbf{O}(\mathbf{r}) \right\|^2, \\ \Psi_j^{\text{opt}}(\mathbf{r}) &= \mathbf{F}^{-1}\Psi_j^{\text{opt}}(\mathbf{u}). \end{aligned} \quad (7)$$

Similarly, the Wirtinger gradient of the $\mathbf{P}(\mathbf{r} - \mathbf{R}_j)$ and $\mathbf{O}(\mathbf{r})$ as well as the

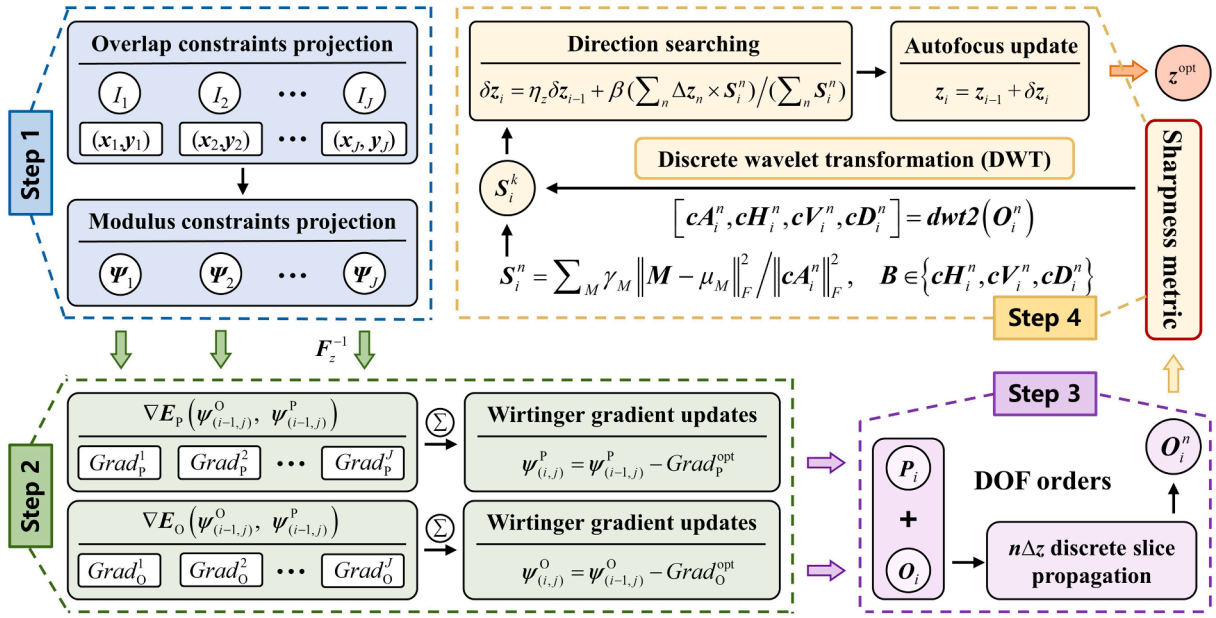


Fig. 1. The detailed workflow of the proposed ptychographic autofocusing algorithm.

gradient descent operation can be expressed as follows

$$\begin{aligned} \nabla E_p &= (\mathbf{P}(\mathbf{r} - \mathbf{R}_j))^* \times (\mathbf{P}(\mathbf{r} - \mathbf{R}_j) \mathbf{O}(\mathbf{r}) - \Psi_j(\mathbf{r})), \mathbf{P}'(\mathbf{r} - \mathbf{R}_j) = \mathbf{P}(\mathbf{r} - \mathbf{R}_j) - \beta_p \nabla E_p, \\ \nabla E_o &= \mathbf{O}^*(\mathbf{r}) \times (\mathbf{P}(\mathbf{r} - \mathbf{R}_j) \mathbf{O}(\mathbf{r}) - \Psi_j(\mathbf{r})), \mathbf{O}'(\mathbf{r}) = \mathbf{O}(\mathbf{r}) - \beta_o \nabla E_o. \end{aligned} \quad (8)$$

To avoid iterative instability, the least squares method is employed to search for the optimal step size [36]. If the update directions are regarded as being orthogonal to each other, the computational overhead is significantly reduced. In this scenario, the non-diagonal entries of the matrix are zero and the optimal step size can be estimated as

$$\begin{pmatrix} \beta_p^{\text{opt}} \\ \beta_o^{\text{opt}} \end{pmatrix} = \begin{pmatrix} \mathbb{R}((\nabla E_p \times \mathbf{O}(\mathbf{r}))^H \times (\Psi_j(\mathbf{r}) - \mathbf{P}(\mathbf{r} - \mathbf{R}_j) \mathbf{O}(\mathbf{r}))) / (\|\nabla E_p \times \mathbf{O}(\mathbf{r})\|_2^2 + \tau) \\ \mathbb{R}((\nabla E_o \times \Psi_j(\mathbf{r}))^H \times (\Psi_j(\mathbf{r}) - \mathbf{P}(\mathbf{r} - \mathbf{R}_j) \mathbf{O}(\mathbf{r}))) / (\|\nabla E_o \times \Psi_j(\mathbf{r})\|_2^2 + \tau) \end{pmatrix}. \quad (9)$$

where, τ is a small regularization constant, and \mathbb{R} is the real number constraint of the step size. After determining the optimal step size, the optimal $\mathbf{P}(\mathbf{r} - \mathbf{R}_j)$ and $\mathbf{O}(\mathbf{r})$ at the j -th scanning position can be reconstructed in Eq. (8). This is the Step 2 in Fig. 1.

2.2. Wavelet-domain autofocusing for specimen-detector distance error

If the specimen-detector distance error occurs, extremely minimal distance deviation dz will lead to perturbations in the ptychographic scanning coordinate, destroying the consistency of the reconstruction algorithm and the physical experiment in terms of the overlapping constraints as,

$$\bar{\mathbf{R}}_j = \left\langle \frac{c_j L}{\lambda(z \pm dz)} \right\rangle \approx \left\langle \mathbf{R}_j \mp \frac{c_j L}{\lambda z^2} dz \right\rangle \quad (10)$$

where, \mathbf{R}_j and $\bar{\mathbf{R}}_j$ are the precise and perturbative scanning coordinate at the j -th scanning position, respectively. c_j is the experimental scanning trajectory of the displacement stage. L is the detector size and λ is the illumination coherent wavelength. z is the specimen-detector distance,

and $\langle \cdot \rangle$ represents a rounding operator. It is obvious to realize that the axial misalignments make the scanning trajectory appear to expand or shrink in the inversion algorithm, which in turn causes reconstruction artifacts due to inconsistencies between the inversion algorithm and physical experiment.

Indeed, scanning trajectories with axial distance errors can be thought of as resulting from propagation from the correct distance to the wrong distance. If the error distance is propagated “backwards” to the correct distance, it is possible to eliminate reconstruction artifacts by obtaining scanning coordinates consistent with physical experiments, and thus reconstructed high-fidelity specimens [29,37]. Therefore, we

performed N discrete slice propagations for the $\mathbf{O}(\mathbf{r})$ estimation specimen of Eq. (8) around the reconstruction plane, and total propagation distance is limited to orders of depth of field (DOF),

$$\begin{aligned} \mathbf{O}(\mathbf{r}, \mathbf{z} + n\Delta\mathbf{z}) &= \mathfrak{S}^{-1} \mathbf{H}_{n\Delta\mathbf{z}}(u_x, u_y) \mathfrak{S} \mathbf{O}(\mathbf{r}, \mathbf{z}), \\ \mathbf{H}_{n\Delta\mathbf{z}}(u_x, u_y) &= \exp \left[i2\pi / \left(\lambda n\Delta\mathbf{z} \sqrt{1 - (u_x)^2 - (u_y)^2} \right) \right]. \end{aligned} \quad (11)$$

where, \mathfrak{S} represents the angular spectrum propagation, and $\Delta\mathbf{z}$ is the propagation interval of discrete slices. In the paraxial approximation, it can be obtained as

$$\Delta\mathbf{z} = \lambda(2\mathbf{z}/L)^2 \quad (12)$$

After discrete slice propagation, the discrete wavelet transforms of the estimated specimen \mathbf{O}_i at n different slice planes can be expressed as

$$[cA_i^n, cH_i^n, cV_i^n, cD_i^n] = \text{dwt2}(\mathbf{O}_i^n), \quad (13)$$

where, cA_i^n , cH_i^n , cV_i^n and cD_i^n are the approximation coefficient, horizontal coefficient, vertical coefficient and diagonal coefficient, respectively. $\text{dwt2}()$ is the two-dimensional discrete wavelet transform, and i is

the iteration number. To enhance the unimodality and minimize the single peak range [38,39], according to the wavelet coefficients, the reconstructed sharpness S_i^n of the specimens from different slice planes are quantified as

$$S_i^n = \sum_M \gamma_M \|\mathbf{M} - \mu_M\|_F^2 / \|\mathbf{cA}_i^n\|_F^2, \quad \mathbf{M} \in \{\mathbf{cH}_i^n, \mathbf{cV}_i^n, \mathbf{cD}_i^n\} \quad (14)$$

where, γ_M is the weighting parameter of the direction coefficients of the matrix \mathbf{M} , and μ_M is the mean value of each element of matrix \mathbf{M} . $\|\bullet\|_F^2$ represents the F -norm operator. By considering the reconstructed sharpness at each sampling distance, the search direction for the axial distance and the corresponding autofocusing update can be expressed as follows

$$\delta \mathbf{z}_i = \eta_z \delta \mathbf{z}_{i-1} + \gamma_z \left(\sum_n n \Delta \mathbf{z} \times S_i^n \right) / \left(\sum_n S_i^n \right), \quad (15)$$

$$\mathbf{z}_i = \mathbf{z}_{i-1} + \delta \mathbf{z}_i.$$

where, $\delta \mathbf{z}_i$ is the axial distance update step of the i -th iteration. η_z and γ_z are the damping factor and feedback factor, respectively. The combination updating of feedback and damping terms is more conducive to accelerate convergence and avoid violent oscillations of the autofocusing algorithm. After completing the update of the axial distance, the ptychographic scanning trajectory in Eq. (10) needs to be updated once again for ptychographic inversion. This is the Steps 3 and 4 in Fig. 1.

By performing steps 1 to 4 until the specimen-detector distance converges and no longer oscillates, the DWT-zPIE autofocusing algorithm will provide the axial distance z , the amplitude-phase reconstruction results of the specimen and the probe. Note that it is preferable to choose alternating iterations between the autofocusing and the position self-calibration [40] to avoid crosstalk between the scanning position errors and the specimen-detector distance error. The detailed workflow of the proposed autofocusing algorithm is shown Algorithm 1 in and Fig. 1.

Input: Measured diffraction field $I_f(\mathbf{u})$, Scanning position coordinates (x_j, y_j)

- 1 **For** $i = 1 : I$ **do** % Iteration numbers
- 2 **For** $j = 1 : J$ **do** % Scanning positions
- 3 **Step 1:** Overlapping and modulus constraints \Rightarrow Eqs. (1-6)
- 4 **Step 2:** Wirtinger gradient updates \Rightarrow Eqs. (7-9)
- 5 **End for**
- 6 **Step 3:** Discrete slice propagation \Rightarrow Eqs. (11-12)
- 7 **Step 4:** Wavelet-domain autofocusing \Rightarrow Eqs. (13-15)
- 8 **End for**

Output: Specimen-detector distance z^{opt} , reconstructed specimen \mathbf{O}^{opt} and probe \mathbf{P}^{opt}

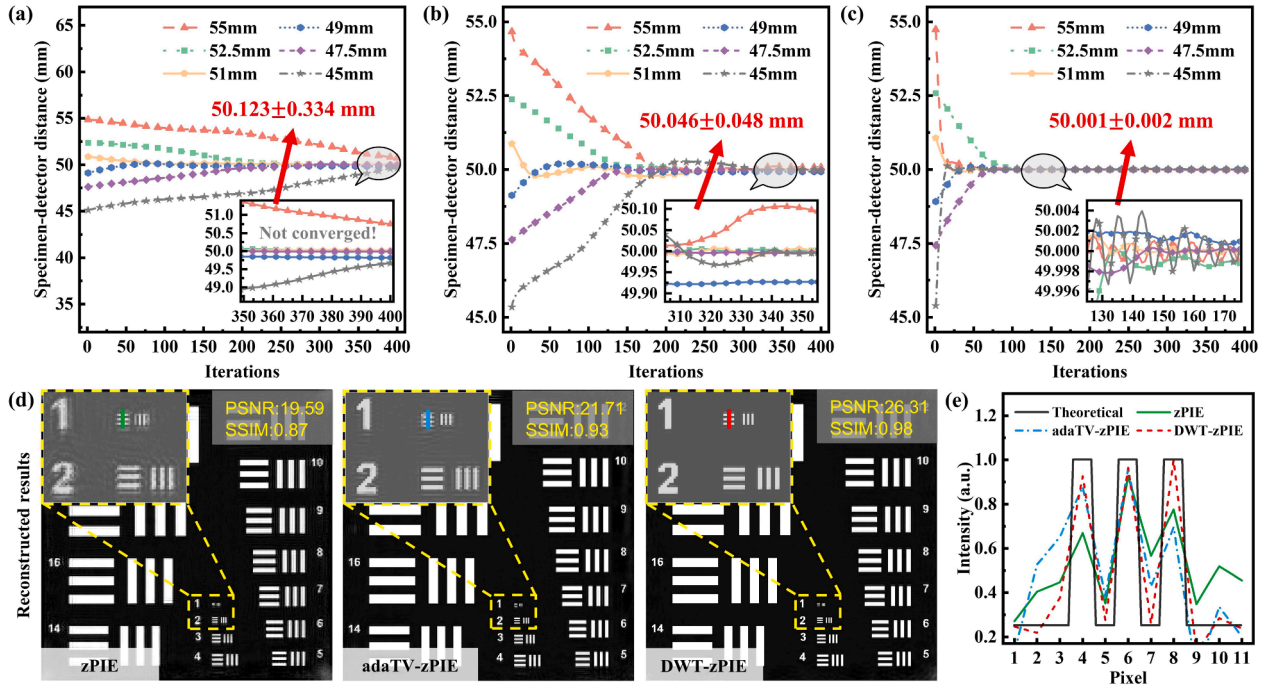


Fig. 2. Numerical simulations: (a) zPIE convergence curve. (b) adaTV-zPIE convergence curve (c) DWT-zPIE convergence curve. (d) Reconstruction results of the varying autofocusing algorithms. (e) Line traces of the region of interest.

3. Numerical simulations

To verify the effectiveness of the proposed algorithm, various autofocusing strategies, including zPIE, adaTV-zPIE, and DWT-zPIE, were compared through numerical simulations. In these simulations, a 632.8 nm coherent beam was focused by a 25 mm lens to form an illuminated probe with a resolution of 256×256 pixels in a $2 \times$ focal plane. The specimen, “resChart,” was mounted at the $2 \times$ focal plane of the lens and driven by a two-dimensional displacement stage for grid scanning. The detector was positioned downstream, 50 mm away from the specimen, to capture coherent diffraction signals at all scanning positions. The scanning step size was set to approximately 24 pixels to ensure a sufficient scanning overlap rate for ptychographic reconstruction. Additionally, an offset of $\sim 10\%$ was applied to the scanning coordinates to eliminate periodic artifacts in the ptychographic reconstruction. To simulate a realistic experimental environment, 35 dB mixed noise [41] was added to the ideal coherent diffraction signals.

To mitigate the influence of non-converged outcomes during the initial iterations on the autofocusing algorithm, a momentum-accelerated ptychographical iterative engine (mPIE) was employed for pre-iterations. After 500 mPIE iterations achieved convergence, various autofocusing strategies were applied to calibrate the axial distance with different initial values. Apart from the differences in autofocusing strategies, the reconstruction parameters and initial conditions remained identical. Fig. 2(a)–(c) illustrate a series of axial distance convergence curves for the different autofocusing algorithms. As shown in Fig. 2(a), the zPIE autofocusing algorithm is highly sensitive to the initial distance. When the true axial distance significantly deviates from the initial value, the zPIE algorithm fails to converge. In contrast, Fig. 2(b) demonstrates that the adaTV-zPIE algorithm, enhanced by the inclusion of the L_p -norm total variation for each pixel, offers a broader initial search range compared to zPIE. Additionally, the adaTV-zPIE algorithm exhibits superior parameter calibration accuracy and convergence uncertainty, even under 35 dB mixed noise conditions. Finally, as shown in Fig. 2(c), the proposed DWT-zPIE autofocusing algorithm achieves the highest performance, with a minimum calibration error of 0.001 mm and convergence uncertainty of 0.002 mm. Compared to the spatial-domain sharpness-based strategies of zPIE and adaTV-zPIE, the DWT-zPIE algorithm integrates multi-dimensional information from both spatial and frequency domains, effectively minimizing noise crosstalk. This integration results in a two-order-of-magnitude improvement in calibration accuracy and convergence uncertainty. A detailed comparison of axial distance calibration results across different autofocusing algorithms is presented in Table 1.

Fig. 2(d) presents the reconstruction results of different autofocusing algorithms. Compared to the zPIE reconstruction results, both the adaTV-zPIE and DWT-zPIE algorithms significantly reduce reconstruction artifacts. Among the smallest line pair features, the DWT-zPIE reconstruction results clearly exhibit the highest clarity, with better

background uniformity. Quantitative image quality assessment metrics, including peak signal-to-noise ratio (PSNR) and structural similarity (SSIM), further validate the superiority of the proposed method over conventional autofocusing algorithms, as shown in Fig. 2(d). Fig. 2(e) illustrates the line trace results for the minimum line pair feature in the region of interest, obtained using different autofocusing algorithms. It is evident that the proposed algorithm outperforms the others in terms of imaging clarity, uniformity, and contrast. In addition to improvements in calibration accuracy and imaging resolution, the proposed algorithm also demonstrates a significant increase in reconstruction speed. Compared to the zPIE and adaTV-zPIE autofocusing algorithms, the proposed method achieves convergence in nearly hundred iterations, resulting in a several-fold improvement in convergence speed.

4. Experiments

To validate the correctness of the numerical results, a prototype system for ptychographic imaging with transmission geometry was constructed. In this transmission geometry, as shown in Fig. 3(a), a 632.8 nm laser (N-STP-912, Newport) passed through filter F1 (NE20A-A, Thorlabs), pinhole P1 (ID25Z, Thorlabs), and mirrors M1 and M2 (BB1-E02, Thorlabs) for beam shaping. The shaped beam was then expanded and collimated by lenses L1 and L2 (GBE05-B, Thorlabs). After focusing through a 50 mm lens L3 (KPX049AR.14, Newport), an illuminated probe was formed. The specimen to be measured was positioned at the back focal plane of lens L3 and was scanned transversely by a two-dimensional displacement stage (L-836.511212, Physik Instrumente). After the illuminated probe interacted with the specimen, the exit waves propagated a distance z to the far-field detector (QHY268M, QHYCCD, pixel size $3.76 \mu\text{m}$) to collect the corresponding diffracted signals. The scanning step size was set to $20 \mu\text{m}$ to ensure more than 80% overlap with the approximately 200 μm beam diameter. In Extend Fullwell 2CMSIT mode, the central 4096×4096 pixel area was cropped to collect 441 raw data points. To enhance computational efficiency, an $8 \times$ binning operation was performed, downsampling to 512×512 pixels for the original diffraction fields. The diffraction signals at different scanning positions are shown in Fig. 3(b).

The first experiment was conducted using the benchmark target USAF-1951 specimen (RES-1, Newport) to verify the effectiveness of the proposed autofocusing algorithm. In this experimental setup, the axial distance between the detector and the benchmark target was approximately 26 mm. After 500 mPIE iterations to eliminate initial blurring and crosstalk, various autofocusing algorithms were applied in ptychographic reconstructions. It is important to note that the ptychographic autofocusing algorithm primarily corrects perturbations in the scanning trajectory caused by deviations in axial distance. Thus, the positioning accuracy of the displacement stage also influences the accuracy of axial distance correction. In numerical simulations, the scanning position is precisely known, whereas in physical experiments, the scanning position is often uncertain. Therefore, scanning position errors are typically addressed using a parallel cross-correlation strategy [40] to reduce calibration crosstalk in system parameters by alternating iterations during the autofocusing process. After 500 autofocusing iterations, the convergence results of different algorithms are shown in Fig. 4(a)–(c). The convergence distance of the zPIE autofocusing algorithm exhibits significant variability with different initial values, as seen in Fig. 4(a). In contrast, the adaTV-zPIE algorithm significantly reduces axial distance calibration uncertainty, as shown in Fig. 4(b). The proposed DWT-zPIE autofocusing algorithm further reduces axial distance uncertainty to 0.02 mm, providing an order of magnitude improvement compared to the zPIE algorithm. Since the real axial distance could not be directly measured in the experiment, the results indicated that the imaging quality of the sample was highest when the axial distance was close to 26.45 mm. Fig. 4(e) displays the reconstruction quality of the benchmark target USAF-1951 at different convergence distances. The adaTV-zPIE algorithm significantly reduces ptychographic reconstruction

Table 1
The comparison results of axial distance calibration.

	Iteration initial value (mm)	Convergence results (mm)		
		zPIE	adaTV-zPIE	DWT-zPIE
Varying initial distance z_0	55.000	50.775	50.124	50.002
	52.500	50.098	50.051	50.001
	51.000	50.122	50.054	50.001
	49.000	49.913	49.974	50.002
	47.500	50.190	50.043	49.998
	45.000	49.767	50.032	50.001
Specimen-detector distance	/	50.123	50.046	50.001
Calibration error	/	0.123	0.046	0.001
Convergence uncertainty	/	0.334	0.048	0.002

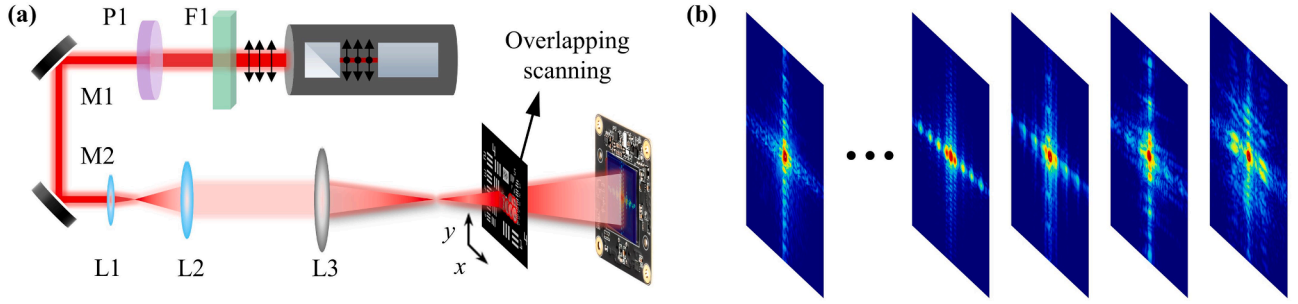


Fig. 3. Experimental setup: (a) Optical schematic. (b) Measured diffraction fields.

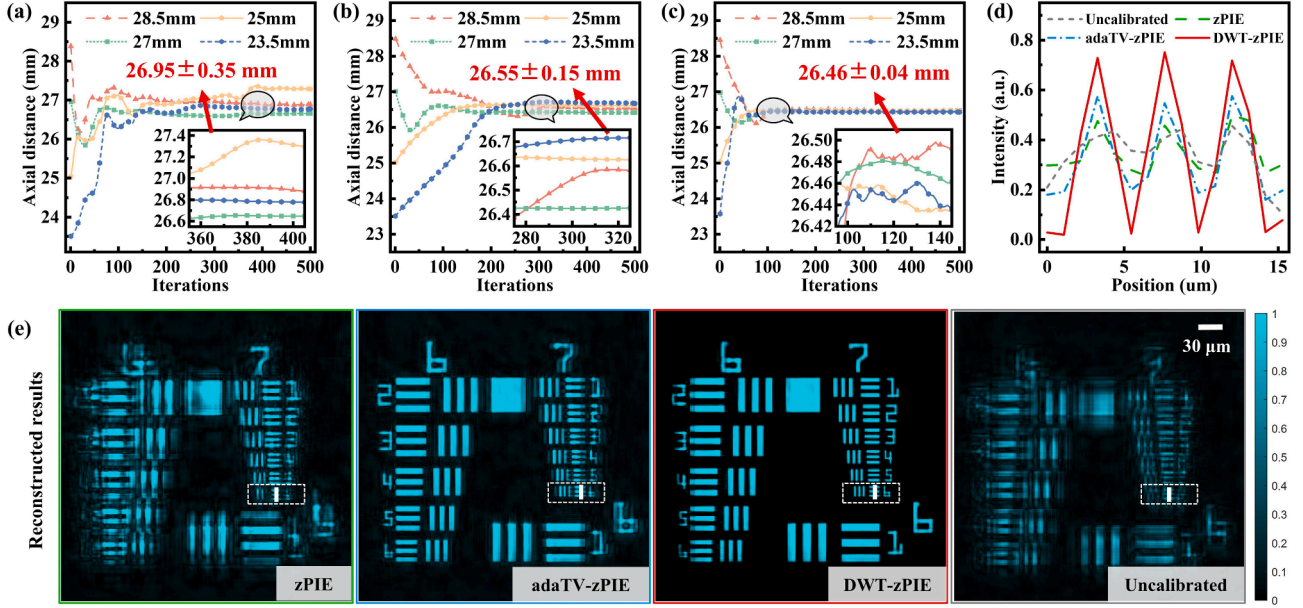


Fig. 4. Experimental results of the USAF-1951: (a) zPIE convergence curve. (b) adaTV-zPIE autofocusing convergence curve (c) DWT-zPIE autofocusing convergence curve. (d) Line traces of the region of interest. (e) Reconstruction results of the varying autofocusing.

artifacts, particularly in the series of line pairs of Group (6), compared to the zPIE algorithm. Among all algorithms, the proposed DWT-zPIE algorithm shows superior reconstruction quality compared to both zPIE and adaTV-zPIE. As observed in numerical simulations, the proposed algorithm demonstrates optimal performance in terms of imaging clarity, uniformity, and contrast. Regardless of the autofocusing algorithm used, the final imaging quality is always superior to the uncalibrated result shown in Fig. 4(e). The quantitative line trace results for the region of interest in Group (7)/Element 6 (full-pitch resolution: $4.384 \mu\text{m}$) further confirm these findings, with the proposed algorithm offering the best clarity and contrast, while the other algorithms suffer from reconstruction artifacts.

The second experiment was conducted with a biological specimen to verify the applicability of the proposed autofocusing algorithm. In this experimental setup, the axial distance was reduced to approximately 14 mm (numerical aperture: 0.482 NA) to further enhance the imaging resolution of ptychography. In the meantime, for the weakly scattering red blood cells, a diffraction signal preprocessing with high dynamic range image fusion [41] was also employed to improve the high-order signal-to-noise ratio. In the convergence results of the adaTV-zPIE algorithms shown in Fig. 5(a) and (b), the biological specimen is significantly more sensitive to initial distance deviations compared to the amplitude-only USAF-1951. In contrast, the proposed DWT-zPIE autofocusing algorithm, as shown in Fig. 5(c), demonstrates strong convergence robustness across various initial distances. According to

uncertainty metrics for axial distance calibration, the proposed algorithm exhibits a several-fold improvement in convergence uncertainty compared to conventional sharpness-based autofocusing algorithms. Fig. 5(e) presents the entire field of view (FOV) reconstruction results of the proposed algorithms at their corresponding convergence distances. In different regions of interest, noticeable differences in imaging fidelity are observed between the zPIE and adaTV-zPIE autofocus algorithms, including both amplitude and phase reconstructions shown in Fig. 5(f) and (g). In contrast, there is virtually no difference in imaging quality across different regions when using the proposed algorithms. The underlying reason for such a phenomenon can be attributed to the fact that zPIE and adaTV-zPIE algorithms often rely on fine iterative parameter optimization and window selection, whereas the proposed algorithm eliminates these constraints and determines the optimal focus across the entire field of view. Meanwhile, the zPIE and adaTV-zPIE algorithms only rely on extracting the grey levels (contrast) of the amplitude-only image. Instead, the algorithm proposed utilizes wavelet transform to process the complex field image (amplitude-phase image), and its multidimensional information extraction greatly reduces the noise crosstalk in the convergence process. However, when the thickness of the biological sample exceeds the “thin sample approximation” in ptychography, the reconstruction of three-dimensional biological samples must rely on slicing modeling to determine the optimal axial distances for different thicknesses [29].

Although the axial distance could not be determined in actual

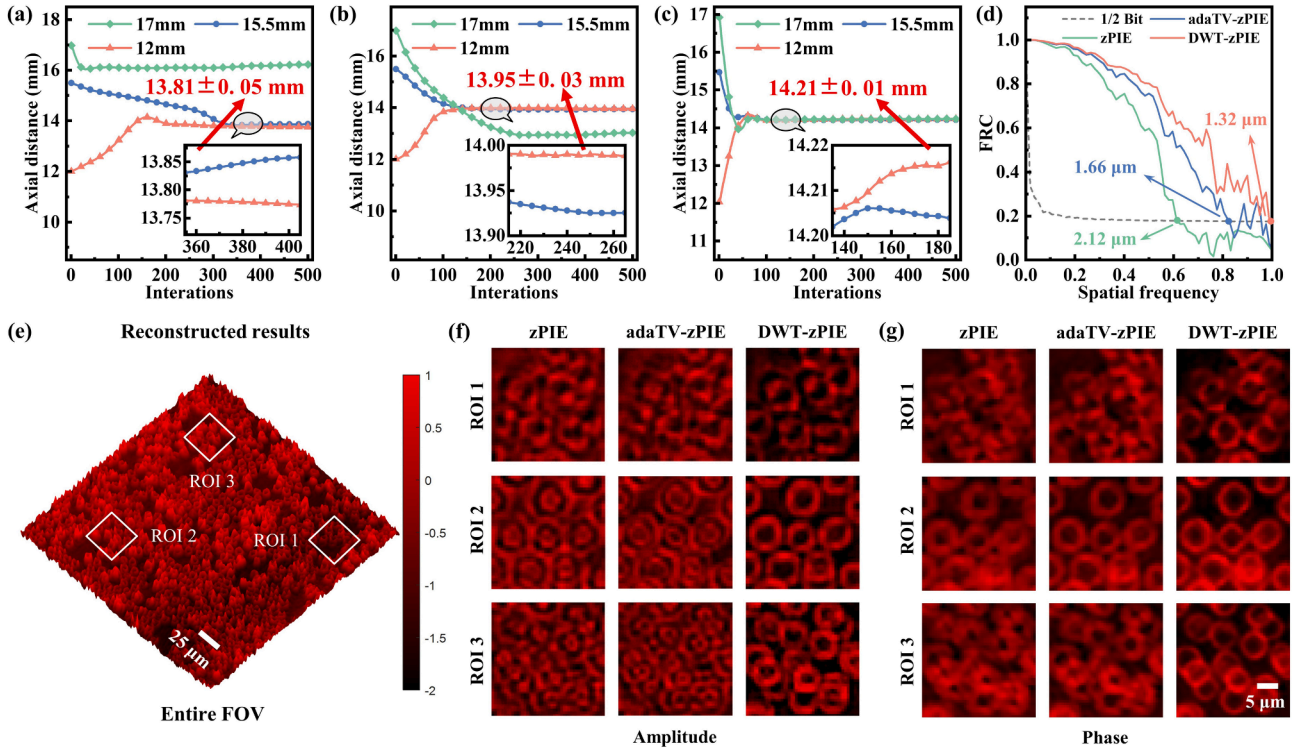


Fig. 5. Experimental results of the biological specimen: (a) zPIE convergence curve. (b) adaTV-zPIE convergence curve (c) DWT-zPIE convergence curve. (d) Results of the Fourier ring correlation. (e) Reconstruction results of DWT-zPIE. (f) and (g) Amplitude-phase results of the regions of interest.

experiments, it is clear that the system will have a higher imaging resolution if the axial distance is fully calibrated. Fig. 5(d) illustrates the reconstructed resolutions using Fourier ring correlation for different autofocus algorithms. After the axial distance calibrating, the zPIE autofocus algorithm achieves a full-pitch resolution of 2.12 μm ptychographic imaging (NA = 0.487). In comparison, the adaTV-zPIE algorithm achieves a full-pitch resolution of 1.66 μm ptychographic imaging, where the NA of the system is 0.483. However, under identical conditions, the proposed DWT-zPIE autofocus algorithm can achieve a 1.32 μm full-pitch resolution in the 0.476NA, delivering nearly a two-fold improvement in imaging resolution.

5. Conclusion

In this paper, we propose an accurate and fast autofocus algorithm for calibrating the specimen-detector axial distance in lensless ptychography imaging. First, we utilize wavelet-domain multidimensional coefficients to quantify reconstruction clarity in virtual planes within the depth of field. Then, the perturbations in the scanning coordinates, which lead to axial distance deviations, are corrected through alternating iterations of the proposed autofocus algorithm and a position self-calibration algorithm. During the iteration process, the proposed algorithm does not require careful window searching or parameter fine-tuning, effectively avoiding local optima, violent convergence oscillations, and crosstalk. Simulations and experiments on both amplitude and complex specimens were performed and compared with conventional sharpness-based autofocus algorithms. The results demonstrate that the proposed algorithm achieves robust convergence in nearly hundred iterations and significantly removes reconstruction artifacts, providing a several-fold to orders of magnitude improvement in convergence speed, calibration accuracy, and uncertainty. More importantly, after axial distance calibration, the proposed algorithm enables the ptychographic imaging system to reach an optimal performance, delivering a nearly two-fold improvement in imaging resolution. These substantial advancements significantly broaden its potential

applications, including in other coherent diffractive imaging techniques such as coded ptychography (coherent modulated imaging), in-line holography, and multi-plane phase retrieval.

In fact, the proposed algorithm typically performs well in scenarios with sufficient exposure. However, when the noise intensity exceeds normal levels, particularly in low-dose, weakly scattering situations, the algorithm tends to suffer from noise crosstalks, which can negatively affect calibration accuracy, convergence speed, and convergence uncertainty, and may even lead to convergence failure. In such cases, potential strategies, such as noise separation or modeling [36,42], could be considered during the reconstruction process in future work.

CRediT authorship contribution statement

Li Liu: Writing – original draft, Validation, Methodology, Data curation, Conceptualization. **Bailin Zhuang:** Software, Formal analysis, Data curation. **Jinxiang Du:** Software, Formal analysis, Data curation. **Lei Zhong:** Visualization, Software, Data curation. **Haoyang Liang:** Software, Resources, Investigation. **Hanqi Chen:** Software, Resources, Investigation. **Qihang Zhang:** Visualization, Software, Resources. **Yunhao Liu:** Visualization, Investigation, Data curation. **Honggang Gu:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Funding acquisition, Conceptualization. **Shiyuan Liu:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Funding acquisition.

Funding

National Natural Science Foundation of China (52130504, 52450258), Key Research and Development Program of Hubei Province (2021BAA013), Guangdong Basic and Applied Basic Research Foundation (2023A1515030149), and Innovation Project of Optics Valley Laboratory (OVL2023PY003).

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.

Acknowledgments

The authors thank the technical support from the Experiment Centre for Advanced Manufacturing and Technology in School of Mechanical Science & Engineering of HUST

Data availability

Data will be made available on request.

References

- [1] H. Pinkard, Z. Phillips, A. Babakhani, D.A. Fletcher, L. Waller, Deep learning for single-shot autofocus microscopy, *Optica* 6 (6) (2019) 794–797.
- [2] J. Du, F. Li, F. Peng, S. Wang, W. Yan, Large-range and high-precision autofocus method based on an annular DOE for a laser direct writing system, *Opt. Express* 30 (5) (2022) 6981–6990.
- [3] Z. Hua, X. Zhang, D. Tu, X. Wang, N. Huang, Learning to high-performance autofocus microscopy with laser illumination, *Measurement* 216 (2023) 112964.
- [4] C.J. Ho, C.C. Chan, H.H. Chen, AF-Net: a convolutional neural network approach to phase detection autofocus, *IEEE Trans. Image Process.* 29 (2020) 6386–6395.
- [5] J. Li, S. Zhang, J. Deng, Y. Liu, Development of a stepping piezoelectric actuator for an autofocus microscopic observation system, *IEEE Trans. Ind. Electron.* 71 (6) (2024) 6107–6116.
- [6] G. Wang, H. Deng, Y. Cai, M. Ma, X. Zhong, X. Gong, Grating-free autofocus for single-pixel microscopic imaging, *Photon. Res.* 12 (6) (2024) 1313–1321.
- [7] Z. Li, R. Zou, W. Kong, X. Wang, Q. Deng, Q. Yan, et al., Terahertz synthetic aperture in-line holography with intensity correction and sparsity autofocusing reconstruction, *Photon. Res.* 7 (12) (2019) 1391–1399.
- [8] G. Dwivedi, L. Pensia, V. Lohchab, R. Kumar, Nondestructive inspection and quantification of soldering defects in PCB using an autofocusing digital holographic camera, *IEEE Trans. Instrum. Meas.* 72 (2023) 7005408.
- [9] C. Guo, Z. Bian, S. Jiang, M. Murphy, J. Zhu, R. Wang, et al., OpenWSI: a low-cost, high-throughput whole slide imaging system via single-frame autofocusing and open-source hardware, *Opt. Lett.* 45 (1) (2020) 260–263.
- [10] Y. Bian, Y. Zhang, P. Yin, H. Li, A. Ozcan, Optical refractometry using lensless holography and autofocusing, *Opt. Express* 26 (23) (2018) 29614–29628.
- [11] Y. Zhang, Z. Huang, S. Jin, L. Cao, Autofocusing of in-line holography based on compressive sensing, *Opt. Lasers Eng.* 146 (2021) 106678.
- [12] X. Zhou, X. Wen, Y. Ji, Y. Geng, S. Liu, Z. Liu, Fast automatic multiple positioning for lensless coherent diffraction imaging, *Opt. Lasers Eng.* 155 (2022) 107055.
- [13] F. Liu, J. Wu, L. Cao, Autofocusing of Fresnel zone aperture lensless imaging for QR code recognition, *Opt. Express* 31 (10) (2023) 15889–15903.
- [14] J. Dohet-Eraly, C. Yourassowsky, F. Dubois, Fast numerical autofocus of multispectral complex fields in digital holographic microscopy with a criterion based on the phase in the Fourier domain, *Opt. Lett.* 41 (17) (2016) 4071–4074.
- [15] H.Y. Chen, W.J. Hwang, C.J. Cheng, X.J. Lai, An FPGA-based autofocusing hardware architecture for digital holography, *IEEE Trans. Comput. Imaging* 5 (2) (2019) 287–300.
- [16] M. Trusiak, J.A. Picazo-Bueno, P. Zdankowski, V. Mico, DarkFocus: numerical autofocusing in digital in-line holographic microscopy using variance of computational dark-field gradient, *Opt. Lasers Eng.* 134 (2020) 106195.
- [17] C. Guo, F. Zhang, X. Liu, Q. Li, S. Zheng, J. Tan, et al., Lensfree auto-focusing imaging using nuclear norm of gradient, *Opt. Lasers Eng.* 222 (156) (2022) 107076.
- [18] Y. Wen, H. Wang, A. Anand, W. Qu, H. Cheng, Z. Dong, et al., A fast autofocus method based on virtual differential optical path in digital holography: theory and applications, *Opt. Lasers Eng.* 121 (2019) 133–142.
- [19] J. Zhang, C. Liu, Y. Yang, Z. Wang, An auto-focusing method for the lens-free single-shot digital holography based on the dissimilar state, *Measurement* 242 (2025) 115912.
- [20] Y. Zhang, H. Wang, Y. Wu, M. Tamamitsu, A. Ozcan, Edge sparsity criterion for robust holographic autofocusing, *Opt. Lett.* 42 (19) (2017) 3824–3827.
- [21] Z. Ren, N. Chen, E.Y. Lam, Automatic focusing for multisectional objects in digital holography using the structure tensor, *Opt. Lett.* 42 (9) (2017) 1720–1723.
- [22] Z. Ren, E.Y. Lam, J. Zhao, Acceleration of autofocusing with improved edge extraction using structure tensor and Schatten norm, *Opt. Express* 28 (10) (2020) 14712–14728.
- [23] S. Jiao, P.W.M. Tsang, T.C. Poon, J.P. Liu, W. Zou, X. Li, Enhanced autofocusing in optical scanning holography based on hologram decomposition, *IEEE Trans. Ind. Inform.* 13 (5) (2017) 2455–2463.
- [24] A.B. Dharmawan, S. Mariana, G. Scholz, P. Hörmann, T. Schulze, K. Triyana, et al., Nonmechanical parfocal and autofocus features based on wave propagation distribution in lensfree holographic microscopy, *Sci. Rep.* 11 (1) (2021) 3213.
- [25] Z. Ren, Z. Xu, E.Y. Lam, Learning-based nonparametric autofocusing for digital holography, *Optica* 5 (4) (2018) 337–344.
- [26] Y. Wu, Y. Rivenson, Y. Zhang, Z. Wei, H. Günaydin, X. Lin, et al., Extended depth-of-field in holographic imaging using deep-learning-based autofocusing and phase recovery, *Optica* 5 (6) (2018) 704–710.
- [27] M. Montoya, M.J. Lopera, A. Gómez-Ramírez, C. Buitrago-Duque, A. Pabón-Vidal, J. Herrera-Ramírez, et al., FocusNET: an autofocusing learning-based model for digital lensless holographic microscopy, *Opt. Lasers Eng.* 165 (2023) 107546.
- [28] L. Huang, T. Liu, X. Yang, Y. Luo, Y. Rivenson, A. Ozcan, Holographic image reconstruction with phase recovery and autofocusing using recurrent neural networks, *ACS Photonics* 8 (6) (2021) 1763–1774.
- [29] L. Loetgering, M. Du, K.S. Eikema, S. Witte, zPIE: an autofocusing algorithm for ptychography, *Opt. Lett.* 45 (7) (2020) 2030–2033.
- [30] R. Ma, D. Yang, T. Yu, T. Li, X. Sun, Y. Zhu, et al., Sharpness-statistics-based autofocusing algorithm for optical ptychography, *Opt. Lasers Eng.* 128 (2020) 106053.
- [31] J. Dou, Z. Gao, J. Ma, C. Yuan, Z. Yang, L. Wang, Iterative autofocusing strategy for axial distance error correction in ptychography, *Opt. Lasers Eng.* 98 (2017) 56–61.
- [32] W. Cao, L. Bai, Y. Xu, C. Kuang, X. Liu, Fast autofocusing strategy for phase retrieval based on statistical gradient optimization, *Opt. Lasers Eng.* 184 (2025) 108636.
- [33] T. Ruan, W. Lv, Y. Tao, J. Zhang, X. Yan, D. Yang, et al., Adaptive total variation based autofocusing strategy in ptychography, *Opt. Lasers Eng.* 158 (2022) 107136.
- [34] Z. Li, X. Zhou, Y. Wang, G. Huang, S. Liu, et al., Flexible and universal autofocus based on amplitude difference of fractional Fourier transform, *Opt. Lasers Eng.* 175 (2024) 107991.
- [35] X. Chang, S. Jiang, Y. Hu, G. Zheng, L. Bian, Pixel super-resolved lensless on-chip sensor with scattering multiplexing, *ACS Photonics* 10 (7) (2023) 2323–2331.
- [36] M. Odstrčil, A. Menzel, M. Guizar-Sicairos, Iterative least-squares solver for generalized maximum likelihood ptychography, *Opt. Express* 26 (3) (2018) 3108–3123.
- [37] L. Lötgering, M. Rose, K. Keskinbora, M. Baluaksian, G. Dogan, U. Sanli, et al., Correction of axial position uncertainty and systematic detector errors in ptychographic diffraction imaging, *Opt. Eng.* 57 (8) (2024) 084106.
- [38] G. Yang, B.J. Nelson, Wavelet-based autofocusing and unsupervised segmentation of microscopic images, in: 2003 IEEE/RSJ IROS, Las Vegas, USA, 2003, pp. 2143–2148.
- [39] H. Xie, W. Rong, L. Sun, Wavelet-Based Focus Measure and 3-D Surface Reconstruction Method for Microscopy Images, in: 2006 IEEE/RSJ IROS, Beijing, China, 2006, pp. 229–234.
- [40] L. Liu, L. Zhong, M. Gong, J. Du, H. Gu, et al., An efficient and robust self-calibration algorithm for translation position errors in ptychography, *IEEE Trans. Instrum. Meas.* 73 (2024) 4503712.
- [41] L. Liu, W. Li, M. Gong, L. Zhong, H. Gu, S. Liu, Resolution-enhanced lensless ptychographic microscope based on maximum-likelihood high-dynamic-range image fusion, *IEEE Trans. Instrum. Meas.* 73 (2024).
- [42] L. Liu, W. Li, L. Zhong, H. Gu, S. Liu, An adaptive noise-blind-separation algorithm for ptychography, *Opt. Lasers Eng.* 169 (2023) 107748.