LASER
& PHOTONICS
REVIEWS

Supporting Information

for *Laser Photonics Rev.*, DOI 10.1002/lpor.202300770

High-Speed Multi-Modal Extended Depth-of-Field Microscopy with an Electrically Tunable Lens

Runnan Zhang, Ning Zhou, Hanci Tang, Minhao Xia, Zewei Cai, Jiasong Sun, Qian Chen and Chao Zuo**

Supporting Information for High-speed multi-modal extended depth-of-field microscopy with an electrically tunable lens

Runnan Zhang[1,2,3], Ning Zhou[1,2,3], Hanci Tang[1,2,3], Minhao Xia[1,2,3], Zewei Cai[1,2,3], Jiasong Sun[1,3,4], Qian Chen[1,3,*], and Chao Zuo[1,2,3,*]

[1]Smart Computational Imaging Laboratory (SCILab), School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing, Jiangsu 210094, China
[2]Smart Computational Imaging Research Institute (SCIRI) of Nanjing University of Science and Technology, Nanjing, Jiangsu 210019, China
[3]Jiangsu Key Laboratory of Spectral Imaging & Intelligent Sense, Nanjing University of Science and Technology, Nanjing, Jiangsu 210094, China
[*]Address all correspondence to Qian Chen, chenqian@njust.edu.cn; Chao Zuo, zuochao@njust.edu.cn

ABSTRACT

This document provides supplementary information for "High-speed multi-modal extended depth-of-field microscopy with an electrically tunable lens".

Contents

Supporting Information S1. Transport of intensity diffraction tomography and partially coherent transfer functions

In the basic theory of optical diffraction tomography (ODT), the physical quantity describing a thick 3D sample is termed as scattering potential $f(\mathbf{r})$, which is a function of the 3D distribution of complex refractive index (RI) $n(\mathbf{r})$ of the object[1]:

$$f(\mathbf{r}) = k_0^2\left[n_0^2 - n^2(\mathbf{r})\right] \tag{S1}$$

where $\mathbf{r} = (x, y, z)$ is the spatial coordinate, $k_0 = 2\pi/\lambda_0$ is the wave vector, with λ_0 being the free-space wavelength, while $n(\mathbf{r}) = n_{\mathrm{Re}}(\mathbf{r}) + in_{\mathrm{Im}}(\mathbf{r})$ and n_0 are the RI of the specimen and its surrounding medium, correspondingly. The real part of the RI, $n_{\mathrm{Re}}(\mathbf{r})$ describes how phase delays accumulate as light passes through the sample, while the imaginary part, $n_{\mathrm{Im}}(\mathbf{r})$, describes absorption. Since the RI of the object is different from that of the surrounding medium, the light propagation and scattering obey the following inhomogeneous wave equation:

$$\left[\nabla^2 + k_m^2\right] U(\mathbf{r}) = -f(\mathbf{r})\, U(\mathbf{r}) \tag{S2}$$

Obviously, the scattering potential outside the sample is 0. In fact, the scattering potential of the sample is equivalent to the complex RI distribution, so $f(\mathbf{r})$ is often called the object function. The total field $U(\mathbf{r})$ can be written as the superposition of the incident field $U_{in}(\mathbf{r})$, and the scattered field $U_s(\mathbf{r})$:

$$U(\mathbf{r}) = U_{in}(\mathbf{r}) + U_s(\mathbf{r}) \tag{S3}$$

while the incident (un-scattered) field satisfies the Helmholtz equation in a homogeneous medium:

$$\left[\nabla^2 + k_m^2\right] U_{in}(\mathbf{r}) = 0 \tag{S4}$$

The scattered field satisfies the following equation:

$$\left(\nabla^2 + k_m^2\right) U_s(\mathbf{r}) = -f(\mathbf{r})\, U(\mathbf{r}) \tag{S5}$$

The partial differential equation expressed in Eq. S5 cannot be directly solved, but based on Green's method, its solution can be expressed as the following integral form:

$$U_s(\mathbf{r}) = \int G\left(\mathbf{r}-\mathbf{r}'\right) f\left(\mathbf{r}'\right) U\left(\mathbf{r}'\right) d\mathbf{r}' = \left[f\left(\mathbf{r}'\right) U\left(\mathbf{r}'\right)\right] \otimes G(\mathbf{r}) \tag{S6}$$

where $G(\mathbf{r})$ is the 3D Green's function corresponding to the Helmholtz equation in a homogeneous medium, which is an outgoing spherical wave originating from the point of observation:

$$U_s(\mathbf{r}) = \int G\left(\mathbf{r}-\mathbf{r}'\right) f\left(\mathbf{r}'\right) U_{in}\left(\mathbf{r}'\right) d\mathbf{r}' + \int G\left(\mathbf{r}-\mathbf{r}'\right) f\left(\mathbf{r}'\right) U_s\left(\mathbf{r}'\right) d\mathbf{r}' \tag{S7}$$

However, if we assume that the light is scattered only once by the object, i.e., the light is deflected only by a single interaction, a linear relation between the resultant first-order scattered field, $U_{s1}(\mathbf{r})$, and the scattering potential of the object can be obtained:

$$U_{s1}(\mathbf{r}) = \int h(\mathbf{r}-\mathbf{r}') f(\mathbf{r}') U_{in}(\mathbf{r}') d\mathbf{r}' \tag{S8}$$

where $h(\mathbf{r})$ is the 3D coherent point spread function (PSF). In order to simplify the derivation, we introduce the first-order Born approximation $U_{s1}(\mathbf{r}) \approx U_s(\mathbf{r})$, the total field can be expressed as:

$$\begin{aligned} U(\mathbf{r}) &= U_{in}(\mathbf{r}) + U_s(\mathbf{r}) \\ &= U_{in}(\mathbf{r}) + \int h(\mathbf{r}-\mathbf{r}') f(\mathbf{r}') U_{in}(\mathbf{r}') d\mathbf{r}' \\ &= \int h(\mathbf{r}-\mathbf{r}') T(\mathbf{r}') U_{in}(\mathbf{r}') d\mathbf{r}' \end{aligned} \tag{S9}$$

where $T(\mathbf{r}) = 1 + f(\mathbf{r}) = 1 + k_0^2\left[n_0^2 - n^2(\mathbf{r})\right]$ is defined as the 3D complex transmittance function of the object. Based on this complex transmittance expression, 3D coherent imaging and 2D coherent imaging can be formally unified. When a 3D object is illuminated by a partially coherent field with the cross-spectral density (CSD) defined as $W_S(\mathbf{r_1}, \mathbf{r_2})$, the CSD of the transmitted field can be expressed as:

$$W_0(\mathbf{r_1}, \mathbf{r_2}) = W_S(\mathbf{r_1}, \mathbf{r_2}) T(\mathbf{r_1}) T^*(\mathbf{r_2}) \tag{S10}$$

According to Van Cittert-Zernike theorem, the CSD of the illumination before reaching the object can be written as:

$$W_S(\mathbf{r_1}, \mathbf{r_2}) = W_S(\mathbf{r_1} - \mathbf{r_2}) = \int S(\mathbf{u}) e^{j2\pi u(\mathbf{r_1}-\mathbf{r_2})} d\mathbf{u} \tag{S11}$$

Note that here $\mathbf{u}$ represents the 3D frequency vector. $S(\mathbf{u})$ refers to the incoherent generalized source defined on the Eward sphere (in the 6f imaging system the source is located at the conjugated plane of the objective pupil):

$$S(\boldsymbol{\mu}, \eta) = S(\boldsymbol{\mu}) \delta\left(\eta - \sqrt{\left(\frac{1}{\lambda}\right)^2 - |\boldsymbol{\mu}|^2}\right) \tag{S12}$$

After 6D Fourier transforms, the transmitted light field needs to multiply the 3D coherent transfer function $H(\mathbf{u})$ twice at the pupil plane, and then the CSD in the image space can be obtained after inverse Fourier transform:

$$W_I(\mathbf{r_1}, \mathbf{r_2}) = \iint \hat{W}_o(\mathbf{u_1}, \mathbf{u_2}) H(\mathbf{u_1}) H^*(\mathbf{u_2}) e^{j2\pi(\mathbf{u_1}\mathbf{r_1} + \mathbf{u_2}\mathbf{r_2})} d\mathbf{u_1} d\mathbf{u_2} \tag{S13}$$

in which $H(\mathbf{u_1})H^*(\mathbf{u_2})$ is the 3D mutual coherent transfer function. Equivalently, it can also be written in the convolution form in the spatial domain:

$$W_I(\mathbf{r_1},\mathbf{r_2}) = \iint W_o(\mathbf{r_1}',\mathbf{r_2}')\,h(\mathbf{r_1}-\mathbf{r_1}')\,h^*(\mathbf{r_2}-\mathbf{r_2}')\,d\mathbf{r_1}'d\mathbf{r_2}' \tag{S14}$$

where $h(\mathbf{r_1})h^*(\mathbf{r_2})$ is the 3D mutual PSF. The image intensity that we can capture is given by the values on the 'diagonal' of the CSD function:

$$I(\mathbf{r}) = W_I(\mathbf{r},\mathbf{r}) = \iint W_S(\mathbf{r_1},\mathbf{r_2})\,T(\mathbf{r_1})\,T^*(\mathbf{r_2})\,h(\mathbf{r}-\mathbf{r_1})\,h^*(\mathbf{r}-\mathbf{r_2})\,d\mathbf{r_1}d\mathbf{r_2} \tag{S15}$$

This formula can be further simplified by substituting Eq. S11 into Eq. S15:

$$I(\mathbf{r}) = \int S(\mathbf{u})\left|\int T(\mathbf{r}')\,h(\mathbf{r}-\mathbf{r}')\,e^{j2\pi\mathbf{u}\mathbf{r}'}d\mathbf{r}'\right|^2 d\mathbf{u} \equiv \int S(\mathbf{u})\,I_{\mathbf{u}}(\mathbf{r})\,d\mathbf{u} \tag{S16}$$

Equation S16 suggests that the final intensity that can be captured at the image plane can be regarded as the intensity (incoherence) superposition of coherent partial images $I_{\mathbf{u}}(r)$ arising from all points over the incoherent generalized source function. Substituting Eq. S11 into Eq. S15 and rewriting it in the form of Fourier integral gives:

$$I(\mathbf{r}) = \iiint S(\mathbf{u})\,\hat{T}(\mathbf{u_1})\,\hat{T}^*(\mathbf{u_2})\,H(\mathbf{u}+\mathbf{u_1})\,H^*(\mathbf{u}+\mathbf{u_2})\,e^{j2\pi r(\mathbf{u_1}-\mathbf{u_2})}d\mathbf{u_1}d\mathbf{u_2}d\mathbf{u} \tag{S17}$$

Separating the contribution of the specimen and imaging system leads to the notion of the 3D transmission cross-coefficient (TCC)[2–4]:

$$TCC(\mathbf{u_1},\mathbf{u_2}) = \int S(\mathbf{u})\,H(\mathbf{u}+\mathbf{u_1})\,H^*(\mathbf{u}+\mathbf{u_2})\,d\mathbf{u} \tag{S18}$$

Noted that the 3D TCC is actually a 6D function. Based on the 3D TCC, Eq. S17 can be written as:

$$I(\mathbf{r}) = \iint \hat{T}(\mathbf{u_1})\,\hat{T}^*(\mathbf{u_2})\,TCC(\mathbf{u_1},\mathbf{u_2})\,e^{j2\pi\mathbf{r}(\mathbf{u_1}-\mathbf{u_2})}d\mathbf{u_1}d\mathbf{u_2} \tag{S19}$$

The above expressions are similar to those derived from 2D partial coherence imaging, except that the dimension of the TCC increases from four to six. It can be calculated from the overlapping region of the generalized source and two shifted generalized apertures (three hemispherical shells). However, the calculation requires a prohibitively large computation time for the 6D integral of the TCC. Considering the real part and the imaginary part of the 3D complex transmission function separately, i.e., the phase component $P(\mathbf{r})$ and the absorption component $A(\mathbf{r})$ (note that the real part of the 3D object function corresponds to the phase component, instead of absorption as in the 2D case):

$$T(\mathbf{r}) = 1 + f(\mathbf{r}) = 1 + k_0^2\left[n_0^2 - n^2(\mathbf{r})\right] = 1 + P(\mathbf{r}) + jA(\mathbf{r}) \tag{S20}$$

Under the first-order Born approximation, the self-interference of the scattered light can be neglected, and the mutual spectrum of the object can be approximated as:

$$\hat{T}(\mathbf{u_1})\hat{T}^*(\mathbf{u_2}) = \delta(\mathbf{u_1})\delta(\mathbf{u_2}) + \delta(\mathbf{u_2})\left[\hat{P}(\mathbf{u_1}) + j\hat{A}(\mathbf{u_1})\right] + \delta(\mathbf{u_1})\left[\hat{P}(\mathbf{u_2}) - j\hat{A}(\mathbf{u_2})\right] \tag{S21}$$

Substituting Eq. S21 into Eq. S19:

$$I(\mathbf{r}) = TCC(\mathbf{0},\mathbf{0}) + 2\,\mathrm{Re}\left\{\int TCC(\mathbf{u},\mathbf{0})\left[\hat{P}(\mathbf{u}) + j\hat{A}(\mathbf{u})\right]e^{j2\pi\mathbf{ru}}d\mathbf{u}\right\} \tag{S22}$$

where $TCC(\mathbf{u},\mathbf{0})$ is the linear part of the TCC, which is the 3D WOTF:

$$WOTF(\mathbf{u}) = TCC(\mathbf{u},\mathbf{0}) = \int S(\mathbf{u}')H(\mathbf{u}'+\mathbf{u})H^*(\mathbf{u}')d\mathbf{u}' \tag{S23}$$

Taking Fourier transform on Eq. S22, and separating the contributions of phase component $P(\mathbf{r})$ and absorption component $A(\mathbf{r})$ gives:

$$\hat{I}(\mathbf{u}) = B\delta(\mathbf{u}) + H_A(\mathbf{u})\hat{A}(\mathbf{u}) + H_P(\mathbf{u})\hat{P}(\mathbf{u}) \tag{S24}$$

where B is the background intensity. It now becomes obvious that the image contrast due to the absorption and phase are decoupled and linear. Because the intensity is always real, the real and even component of $WOTF(\mathbf{u})$ corresponds to the partially coherent 3D AOTF $H_A(\mathbf{u})$, and the imaginary and odd component of $WOTF(\mathbf{u})$ corresponds to the 3D POTF $H_P(\mathbf{u})$ of the object[5–7]:

$$H_A(\mathbf{u}) = WOTF(\mathbf{u}) + WOTF^*(-\mathbf{u}) \tag{S25}$$

$$H_P(\mathbf{u}) = WOTF(\mathbf{u}) - WOTF^*(-\mathbf{u}) \tag{S26}$$

Supporting Information S2. OTF analysis for different illumination schemes

As described in the manuscript, the annular illumination scheme is adopted. First, we further explain the importance of matched illumination condition from the scanning absorption optical transfer function (sAOTF). As shown in Figs. S1a1-a4, the cut-off frequency of sAOTF is highly related to the numerical aperture of illumination, i.e., the coherent parameter S. Figure S1a5 clearly depicts the 1D line profiles of different sAOTF under different S. We can draw conclusions that the matched illumination condition is critical for spatial resolution, i.e., up to the incoherent diffraction limit. Moreover, the annulus width will influence the imaging contrast, which can be deduced from the transfer functions. As discussed in the manuscript, we adopted annular illumination rather than a circular illumination scheme for enhancing the imaging contrast, while the annular width should be chosen appropriately according to the sAOTF. As shown in Figs. S1b1-b4, if the annulus is too wide, the imaging contrast cannot be strongly boosted, while if the annulus is too narrow, the transfer function is abrupt at certain high frequencies. Therefore, we have to choose the appropriate annulus width experimentally, and we have chosen $\Delta S = 0.2$ in our previous experiments.

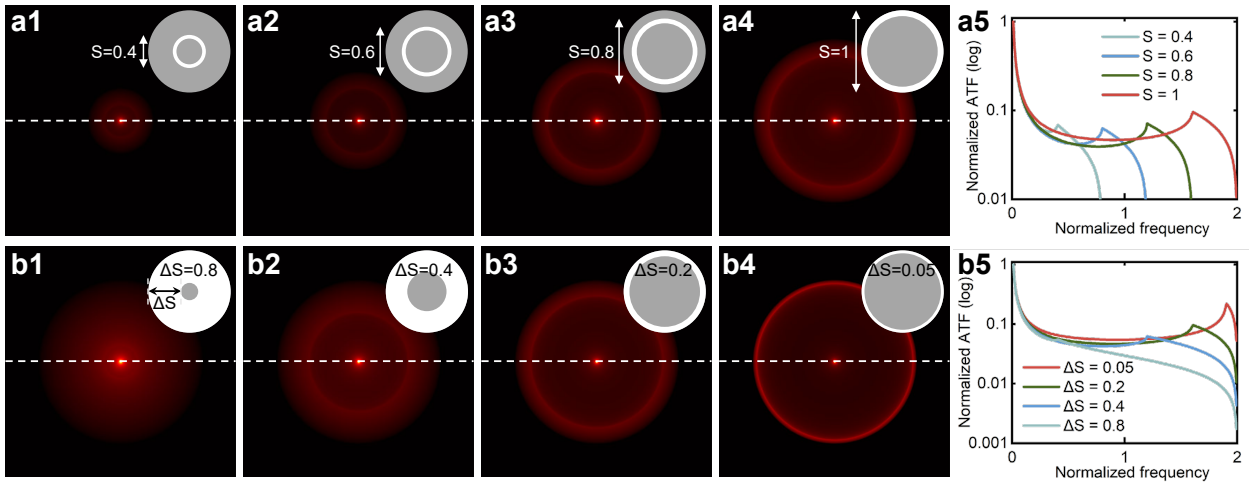

Figure S1. Annular illumination scheme under different numerical aperture and annulus width. a1-a4, The sAOTFs are shown respectively under $S = 0.4, 0.6, 0.8$ and 1 when the source is a narrow annulus. a5, The corresponding 1D line profiles under different S. b1-b4, The sAOTFs are shown respectively under matched illumination conditions with different annulus widths $\Delta S = 0.8, 0.4, 0.2, 0.05$. b5, The corresponding 1D line profiles under different annulus widths. The spatial frequency coordinate is normalized against the coherent resolution limit NA_{obj}/λ. The normalized sAOTF value is displayed in log format.

Supporting Information S3. Relation to the Fourier Slice Theorem

The classical version of the Fourier Slice Theorem states that the 1D slice of a 2D function's Fourier spectrum is the Fourier transform of an orthographic integral projection of the 2D function. The projection and slicing geometry are illustrated in Fig. S2a. This is the basic principle of the conventional computed tomography (CT). We first consider a 2D function $f(x,y)$, and its projection can be represented as the integration along a straight line $l(s,\theta)$:

$$R_f(s,\theta) = \iint f(x,y)\,\delta(x\cos\theta + y\sin\theta - s)\,dxdy \tag{S27}$$

The line equation $l(s,\theta)$ can be expressed as:

$$x\cos\theta + y\sin\theta = s \tag{S28}$$

where s is the distance from the origin to the line, and θ stands for the positive angle from the line to the y-axis (or the positive angle from the projection plane to the x-axis). When θ and s are fixed, $l(s,\theta)$ represents a straight line with the angle θ between the y-axis and the distance s from the origin. The projection of $f(x,y)$ onto the line can be obtained by Eq. S27.

As the central theorem in classical CT, the Fourier slice theorem connects the projection data to the original function in Fourier space, which allows efficient CT reconstructions with FFT. Based on the definition of Fourier transform and variable substitution, it can be proven that the 1D Fourier transform $\hat{R}_f(\gamma,\theta)$ (γ corresponding to the spatial frequency s coordinate) of the 1D projection at angle θ is distributed along a straight line at the same angle in the Fourier space of the 2D object $f(x,y)$.

$$\hat{R}_f(\gamma,\theta) = \hat{f}(u,v)\Big|_{v=\gamma\sin\theta}^{u=\gamma\cos\theta} = \hat{f}(\gamma\cos\theta, \gamma\sin\theta) \tag{S29}$$

where (γ,θ) is the polar coordinates in the frequency domain.

We extend the 2D Fourier Slice Theorem to 3D. The 3D point spread function and optical transfer function are Fourier transform pairs. The 2D slice of a 3D OTF is the 2D Fourier transform of an orthographic integral projection of the 3D PSF. In this paper, the scanning 2D PSF is a vertical projection of 3D PSF, and its corresponding 2D Fourier spectrum is the middle slice of 3D OTF, as shown in Fig. S2b.

Moreover, the Fourier Slice Theorem is especially essential for fluorescent imaging. Traditional CT techniques usually use coherent light sources (such as lasers) to illuminate the object to obtain a specific direction of projection under transmissive imaging, and the projection intensity is a slice of the Fourier spectrum. The ODT techniques take into account diffraction effects, where the projection extends from a slice in the spectrum to a sphere. In order to recover complete 3D refractive index, it is usually necessary to use coherent light sources. The object and illumination beam must relatively rotate to each other

to collect multi-directional projections, and 3D reconstruction can only be performed on samples with absorption or scattering characteristics. However, fluorescent imaging methods are mainly based on the principles of axial scanning and optical sectioning to reconstruct 3D objects, and there is currently a lack of methods to explain 3D reconstruction from the perspective of CT. Our recently published work[8] has found that using an annular aperture can modulate the PSF of the imaging system into a Bessel beam, keeping the PSF invariable in the axial direction and achieving the EDOF imaging. The incoherent Fourier Slice Theorem is derived, and it has found that the 2D intensity signal detected under the annular aperture is equal to a 2D slice in the 3D object spectrum, revealing that the annular aperture modulation can apply CT technique to incoherent fluorescent wide-field microscopy. The use of a fast-scanning ETL for EDOF imaging can be combined with the PALFM technique in order to replace the annular aperture, which is a promising work in the future.

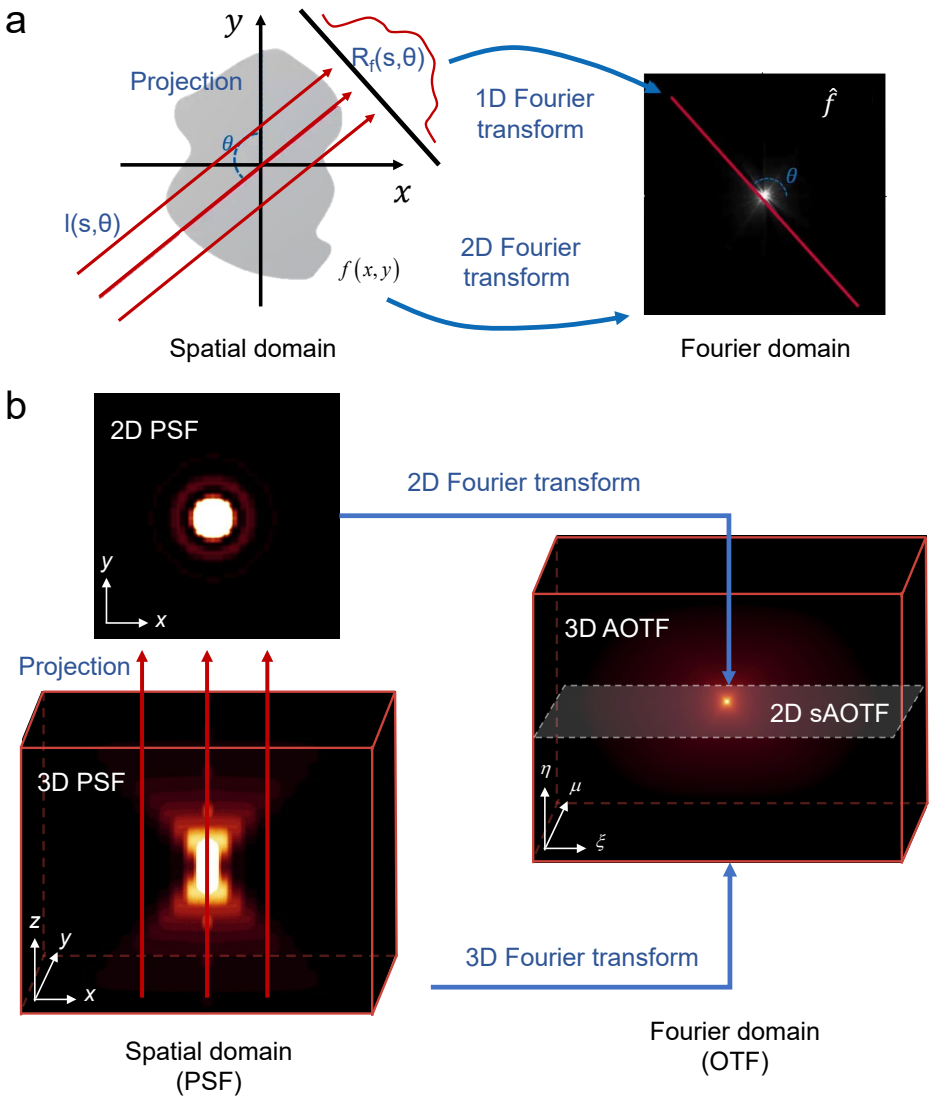

Figure S2. Generalization of the Fourier Slice Theorem. a, classical Fourier Slice Theorem, the 1D slice of a 2D function's Fourier spectrum is the Fourier transform of an orthographic integral projection of the 2D function. b, generalized Fourier slice theorem for OTF, the 2D slice of a 3D OTF is the 2D Fourier transform of an orthographic integral projection of the 3D PSF.

Supporting Information S4. Telecentric defocusing with an ETL

Within the scope of scalar diffraction, free-space propagation is described by means of a linear shift-invariant system. Given the complex amplitude $u_0(x,y)$ of a wave field across the plane $z=0$, the complex amplitude $u_z(x,y)$ across a parallel plane in a distance z is obtained from the convolution with the impulse response of propagation[9]:

$$u_z(x,y) = u_0(x,y) \otimes h_z(x,y) \tag{S30}$$

According to the convolution theorem, Eq. S30 turns into a modulation in the Fourier domain (u,v) and yields:

$$u(u,v) = u_0(u,v) \cdot H_z(u,v) \tag{S31}$$

where the $H_z(u,v)$ is a Fourier transform of $h_z(x,y)$. The corresponding transfer function is a phasor with unit amplitude:

$$H_z(u,v) = \exp\left[-i\pi\lambda z\left(u^2+v^2\right)\right] \tag{S32}$$

where $\lambda = \lambda_0/n$, λ_0 is the wavelength and n is the refractive index of the immersion medium. According to the manuscript, the 4f system presented in Fig. 3 may be used to generate propagated representations $u_z(x,y)$ of the initial wave field $u_0(x,y)$ across the sensor plane, where the index z refers to the associated propagation distance along the optical axis. In the following, the corresponding complex transmittance $t_l(\mu,\xi)$ to be generated by the electrically tunable lens (ETL) shall be investigated. The analysis is based on the transfer function given by Eq. S32 and the geometric parameters of the setup, such as the focal length of the lenses. Considering the Fourier transform property of the first lens, the complex amplitude of the wave field across the (μ,ξ) plane is proportional to the Fourier transform of the complex amplitude incident in the (u,v) plane:

$$u(\mu,\xi) = \frac{1}{i\lambda f} \cdot \mathscr{F}\{u(x,y)\}\left(\frac{\mu}{\lambda f}, \frac{\xi}{\lambda f}\right) \tag{S33}$$

The operator $\mathscr{F}$ denotes the Fourier transform and f is the focal length of the Fourier lenses. From Eq. S33, the substitution $(u,v) = \left(\mu\cdot(\lambda f)^{-1}, \xi\cdot(\lambda f)^{-1}\right)$ is obtained.

Considering the Fourier transform property of the 4f system, the variation of ETL's focal power located at the Fourier plane is equivalent to adding a quadratic phase modulation[10–12]:

$$t_l(\mu,\xi) = \exp\left[-\frac{i\pi}{\lambda f_{eff}}\left(\mu^2+\xi^2\right)\right] \tag{S34}$$

where (μ,ξ) is the spatial coordinates in the Fourier plane of the 4f system, f_{eff} is the effective focal

length of ETL controlled by current.

$$f_{eff} = \frac{f_{ETL} f_{OL}}{f_{ETL} + f_{OL} - d} \tag{S35}$$

where f_{OL} is the focal length of offset lens (OL), d is the axial distance between the two lenses. Note here we consider the composite lens as 'thin' for simplicity, because d is very small and f_{eff} is much larger than the thickness of the two lenses. Comparing Eq. S20 with Eq. S32, we can see the $t_l(\mu, \xi)$ share the same format as the propagation transfer function. With the use of the connection between the frequency coordinates in the Fourier domain and the spatial coordinates in the Fourier plane of the 4f setup $(u, v) = \left(\mu \cdot (\lambda f)^{-1}, \xi \cdot (\lambda f)^{-1}\right)$, the relationship between the focal length variations versus the propagation distances can be established:

$$\Delta z = -\frac{1}{M^2} \frac{n \cdot f^2}{f_{eff}} \tag{S36}$$

where M is the magnification of the objective lens, and f is the focal length of two Fourier lenses. Since f_{eff} can be electronically controlled, the defocus distance Δz can be easily adjusted.

Generally, the scan range is proportional to the square of the focal length of the relay lenses f. However, there is an upper limit to f. To maintain the full NA of the detection lens, the ratio of the focal lengths of the relay lens f and the tube lens f_{TL} must not be larger than the ratio of the aperture of the ETL d_{ETL} and the diameter of the BFP of the detection lens d_{BFP}, i.e.,

$$f \leqslant f_{TL} \cdot \frac{d_{ETL}}{d_{BFP}} \tag{S37}$$

Supporting Information S5. Richardson-Lucy deconvolution with the total variation (TV) regularization

Although the rapid focus scanning collects all focus information, the captured single-shot image is still degraded by out-of-focus light of each layer and by Poisson noise due to photon-limited detection, especially under rapid exposure conditions. While the deconvolution method tends to amplify noise, the usage of image priors-based regularization constraints has to be applied to reduce the noise sensitivity. Here, we combine the iterative Richardson-Lucy deconvolution algorithm[13–15], which computes maximum likelihood estimation adapted to Poisson statistics, with a regularization constraint based on Total Variation (RL-TV) to suppress unstable oscillations while preserving object edges.

Light emission is well-approximated by a Poisson process, and so is the 3D image. If an object o is observed as an image i through an optical system with a PSF h and degraded by a Poisson noise, we can write the likelihood as:

$$p(i|o) = \prod_{\mathbf{x}} \left(\frac{\left[(h \otimes o)(\mathbf{x})^{i(\mathbf{x})} e^{-(h \otimes o)(\mathbf{x})} \right]}{i(\mathbf{x})!} \right) \tag{S38}$$

Maximizing this probability with respect to o and using a multiplicative gradient-based algorithm leads to the RL algorithm:

$$o_{k+1}(\mathbf{x}) = \left\{ \left[\frac{i(\mathbf{x})}{(o_k \otimes h)(\mathbf{x})} \otimes h(-\mathbf{x}) \right] \right\} \times o_k(\mathbf{x}) \tag{S39}$$

The non-regularized RL algorithm minimizes the functional $-\log p(i|o) = J_1(o)$ (up to an additive constant), giving the maximum likelihood estimation:

$$J_1(o) = \sum_x \left(-i(\mathbf{x}) \log \left[(o \otimes h)(\mathbf{x}) \right] + (o \otimes h)(\mathbf{x}) \right) \tag{S40}$$

We introduce a priori constraint on the object by adding to the energy J_1, a regularization term J_{reg}, defined by the TV of the solution o,

$$J_{reg}(o) = \varepsilon_{TV} \sum_s |\nabla o(\mathbf{x})| \tag{S41}$$

The total functional to be minimized is then:

$$J_1(o) + J_{reg}(o) = \sum_{\mathbf{x}} \left(-i(\mathbf{x}) \log \left[(o \otimes h)(\mathbf{x}) \right] + (o \otimes h)(\mathbf{x}) \right) + \varepsilon_{TV} \sum_{\mathbf{x}} |\nabla o(\mathbf{x})| \tag{S42}$$

Using the $L1$ norm over ∇o rather than the $L2$ norm allows for the recovery of a smooth solution with sharp edges. It can be shown that the smoothing process introduced by J_{reg} acts only in the direction tangential to the image level lines and not in the orthogonal direction, so that edges are preserved. The

derivative of J_{reg} with respect to o is a nonlinear term $\frac{\partial}{\partial o}J_{reg} = -\varepsilon_{TV} div\left(\frac{\nabla_{o_k}(\mathbf{x})}{|\nabla_{o_k}(\mathbf{x})|}\right)$. We minimize Eq. S42 using the multiplicative gradient-based algorithm, and we adopt an explicit scheme, defined by:

$$o_{k+1}(\mathbf{x}) = \left\{\left[\frac{i(\mathbf{x})}{(o_k \otimes h)(\mathbf{x})} \otimes h(-\mathbf{x})\right]\right\} \times \frac{o_k(\mathbf{x})}{1 - \varepsilon_{TV} div\left(\frac{\nabla_{o_k}(\mathbf{x})}{|\nabla_{o_k}(\mathbf{x})|}\right)} \tag{S43}$$

where $div(\cdot)$ denotes the divergence, $\nabla_{o_k}(\mathbf{x})$ stands for the gradient of $o_k(\mathbf{x})$, ε_{TV} represents the regularization parameter, and $h(\mathbf{x})$ is the PSF of the optical system. Numerically, we noticed that the regularization parameter ε_{TV} should be neither too small nor too large: if ε_{TV} is too small, RL will be dominated by the data model; if ε_{TV} is too large, RL will be dominated by the regularization term. For larger ε_{TV}, the denominator of Eq. S43 will become zero or negative. This must be avoided because small denominators create points of very high intensity, which are amplified at each iteration. A negative value violates the nonnegativity constraint.

To quantify the quality of the deconvolution we use the I-divergence[16] criteria. According to Ref.[16], the I-divergence is the only consistent measure in the presence of a non-negativity constraint. The I-divergence between two images A and B is given by:

$$I_{A,B} = \sum_{i,j}\left\{A_{i,j}\ln\left[\frac{A_{i,j}}{B_{i,j}}\right] - \left(A_{i,j} - B_{i,j}\right)\right\} \tag{S44}$$

while the mean square error is given by:

$$MSE_{A,B} = \sum_{i,j}\left\{A_{i,j} - B_{i,j}\right\}^2 \tag{S45}$$

We can notice that the I-divergence is non-symmetric, but MSE is. For simulated data, the I-divergence and MSE values can be computed by comparing the final estimate to the initial undegraded object. However, we have problems with real data because we do not know exactly what the initial object before degradation was. In this case, to stop the iterations, we define a difference measure between two successive iterations. If the difference is smaller than a threshold, we stop the computation, and we suppose that the last estimation is the best one. The criterion is defined as the following:

$$\chi_{k+1} = \frac{\sum_{x,y=1}^{N_{x,y}} |o_{k+1}(x,y) - o_k(x,y)|}{\sum_{x,y=1}^{N_{x,y}} o_k(x,y)} < t \tag{S46}$$

Typical threshold values t used are around 10^{-4}, depending on the precision we want. It also depends on the regularization parameter. For high values of the regularization parameter, the stop criterion should not be too small because we noticed that the χ_k could oscillate between two values, as shown in Fig. S3.

Figure S3 shows deconvolution results under different iteration times and regularization parameters.

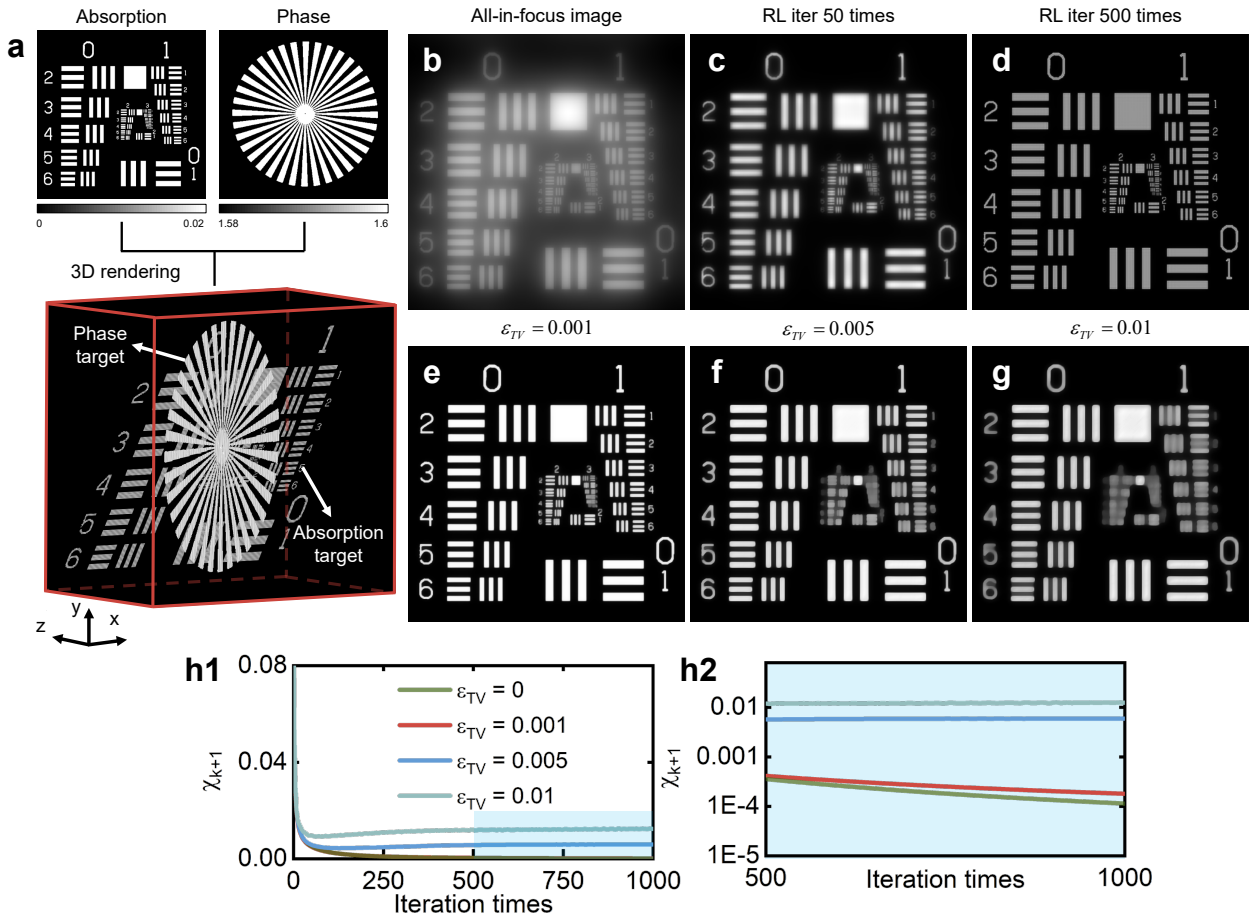

Figure S3. Deconvolution under different iterations and regularization parameters. a, the 3D refraction index. b, single-shot scanning image. c, the RL deconvolution result under 50 iterations. d, the RL deconvolution result under 500 iterations. e-g, RL+TV deconvolution results under different regularization parameters ε_{TV} respectively with 500 iterations. h1, the criterion lines under different iterations. h2, the enlarged region in h1.

We have simulated a 3D object composed of absorption (a tilted USAF target) and phase (a tilted star target). The all-in-focus single-shot scanning image is shown in Fig. S3b. We perform RL deconvolution and judge the stop criteria. The result performed under 500 iterations (Fig. S3d) is much clearer than that under 50 iterations (Fig. S3c). Moreover, the deconvolution results under different regularization parameters ε_{TV} are shown in Figs. S3e-g. It can be seen that the value should be selected appropriately. High values of the regularization parameter will result in over-smoothed results, and the χ_k will oscillate as shown in Figs. S3h1-h2.

Supporting Information S6. Comparisons between RL, RL-TV and Wiener deconvolution under different noise levels.

In this section, we further analyze different deconvolution methods, including RL, RL-TV and Wiener deconvolution, with the influence of different signal-to-noise (SNR) levels and different regularization parameters. The SNR is calculated as:

$$SNR = 10\log_{10}\left[\frac{\sum_{x=1}^{N}\sum_{y=1}^{M} f(x,y)^2}{\sum_{x=1}^{N}\sum_{y=1}^{M}\left[f(x,y)-g(x,y)\right]^2}\right] \tag{S47}$$

where $f(x,y)$ is the original all-in-focus image, and $g(x,y)$ is the noisy image. Figure S4 demonstrates the simulation results. As shown in Fig. S4a1, the all-in-focus image is blurred without noise, and the deconvolution results under different regularization parameters $\varepsilon_{TV} = 0, 0.001, 0.01$ are shown respectively in Figs. S4a2-a4. In Figs. S4b1, c1, d1, the blurred image is added with random noise, and the SNR levels decrease, i.e., 70dB, 51dB, and 20dB respectively. As we can see, under a low SNR level, a low value of the regularization parameter ε_{TV} will result in the deconvolution being ruined by noise while a high value of ε_{TV} will lead to a smooth result.

In addition to the RL-TV deconvolution method, we also make a comparison with Wiener deconvolution, which is a direct deconvolution method calculated by:

$$F(u,v) = \frac{H^*(u,v)\,G(u,v)}{|H(u,v)|^2 + \varepsilon} \tag{S48}$$

where $F(u,v)$ is the Fourier spectrum of object, $G(u,v)$ is the Fourier spectrum of the captured image, $H(u,v)$ is the transfer function and ε is the regularization parameter.

Compared with RL-TV deconvolution, the speed of Wiener deconvolution is fast. However, as can be seen in Figs. S4a5-d5, this deconvolution method is very sensitive to noise. When the noise level is high, the deconvolution result will be corrupted by the noise. Taking both detail perseveration and noise suppression into consideration, we choose RL-TV deconvolution method in all experiments, and the regularization parameters are empirically chosen between 0 and 0.01 according to noise levels.

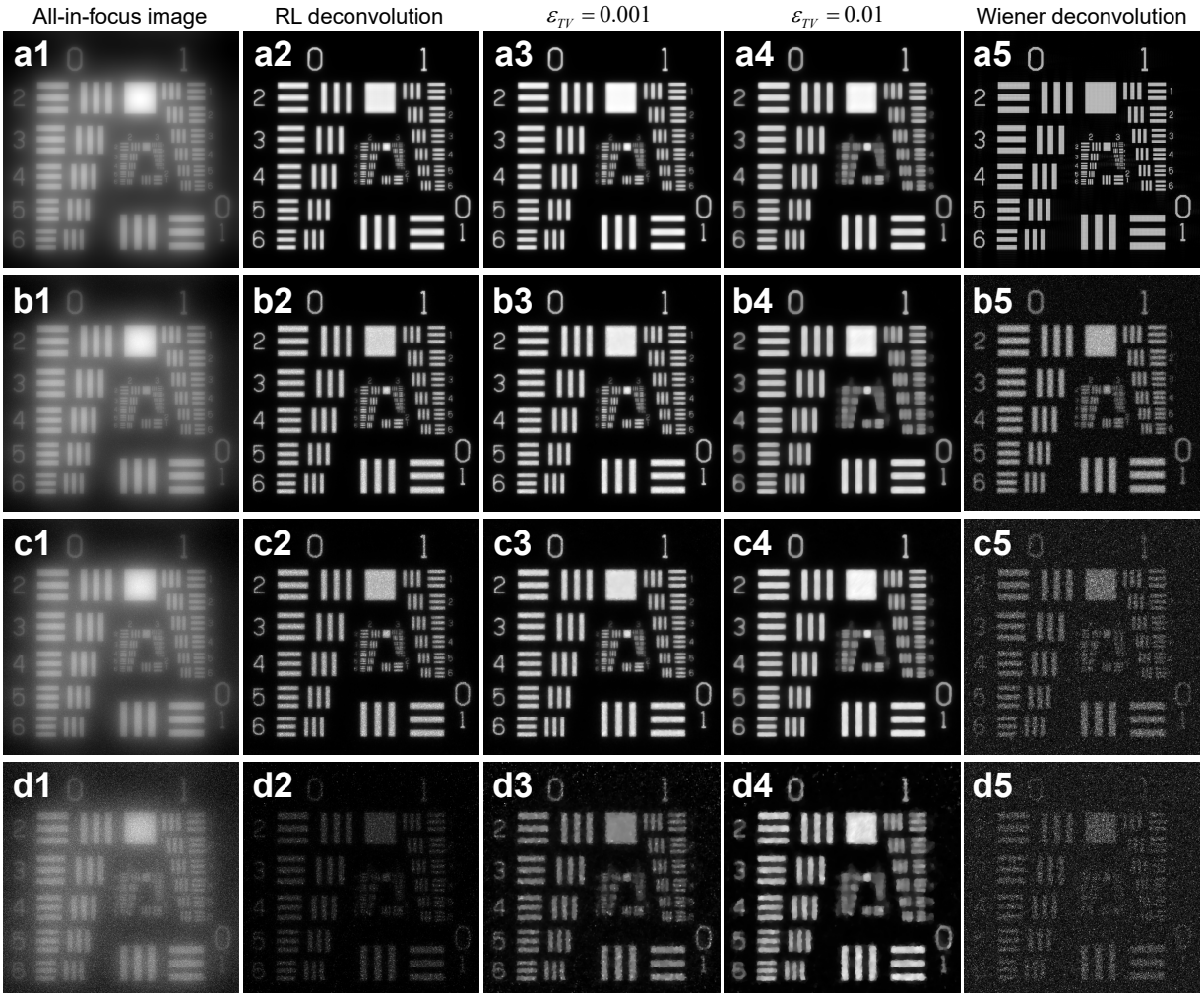

Figure S4. Comparisons between RL deconvolution, RL deconvolution with different regularization parameters and Wiener deconvolution under different SNR levels. Deconvolution results of different SNR levels. a1-d1, the all-in-focus image with different SNR levels, SNR = ∞, 70dB, 51dB, 20dB respectively. a2-d2, RL deconvolution results without TV regularization. a3-d3, RL deconvolution with $\varepsilon_{TV} = 0.001$. a4-d4, RL deconvolution with TV parameter $\varepsilon_{TV} = 0.01$. a5-d5, Wiener deconvolution.

References

1. Wolf, E. Three-dimensional structure determination of semi-transparent objects from holographic data. *Opt. Commun.* **1**, 153–156 (1969).
2. Sheppard, C. & Choudhury, A. Image formation in the scanning microscope. *Int. J. Opt.* **24**, 1051–1073 (1977).
3. Hopkins, H. H. On the diffraction theory of optical images. *Proc. Math. Phys. Eng. Sci.* **217**, 408–432 (1953).
4. Zuo, C. *et al.* High-resolution transport-of-intensity quantitative phase microscopy with annular illumination. *Sci. Rep.* **7**, 7654 (2017).
5. Zuo, C. *et al.* Transport of intensity equation: a tutorial. *Opt. Lasers Eng.* **135**, 106187 (2020).
6. Li, J. *et al.* Transport of intensity diffraction tomography with non-interferometric synthetic aperture for three-dimensional label-free microscopy. *Light. Sci. Appl.* **11**, 154 (2022).
7. Sun, J., Zuo, C., Zhang, J., Fan, Y. & Chen, Q. High-speed fourier ptychographic microscopy based on programmable annular illuminations. *Sci. Rep.* **8**, 7669 (2018).
8. Cai, Z., Zhang, R., Zhou, N., Chen, Q. & Zuo, C. Programmable aperture light-field microscopy. *Laser Photonics Rev.* 2300217 (2023).
9. Goodman, J. W. *Introduction to Fourier optics* (Roberts and Company publishers, 2005).
10. Zuo, C., Chen, Q., Qu, W. & Asundi, A. High-speed transport-of-intensity phase microscopy with an electrically tunable lens. *Opt. Express* **21**, 24060–24075 (2013).
11. Fahrbach, F. O., Voigt, F. F., Schmid, B., Helmchen, F. & Huisken, J. Rapid 3d light-sheet microscopy with a tunable lens. *Opt. Express* **21**, 21010–21026 (2013).
12. Grewe, B. F., Voigt, F. F., vant Hoff, M. & Helmchen, F. Fast two-layer two-photon imaging of neuronal cell populations using an electrically tunable lens. *Biomed. Opt. Express* **2**, 2035–2046 (2011).
13. Lucy, L. B. An iterative technique for the rectification of observed distributions. *Astron. J.* **79**, 745 (1974).
14. Richardson, W. H. Bayesian-based iterative method of image restoration. *J. Opt. Soc. Am.* **62**, 55–59 (1972).
15. Dey, N. *et al.* Richardson–lucy algorithm with total variation regularization for 3d confocal microscope deconvolution. *Microsc. Res. Tech.* **69**, 260–266 (2006).
16. Csiszar, I. Why least squares and maximum entropy? an axiomatic approach to inference for linear inverse problems. *Ann. Stat.* 2032–2066 (1991).