Supplementary Information for Transport-of-intensity diffraction tomography with nonuniform axial sampling (TIDT-NAS)

Yefeng Shu[1,2,3,†], Jiasong Sun[1,2,3,4,†], Shun Zhou[1,2,3], Yao Fan[1,2,3,4], Anqi Tang[1,2,3], Wenzhuo Xie[1,2,3], Qingyang Fu[1,2,3], Mohammad Reza Zarei[1,2,3], Xiyu Chen[1,2,3], Xiaolei Zhang[2,5], Qian Chen[1,2,3,4,5,*], and Chao Zuo[1,2,3,4,5,*]

[1]Smart Computational Imaging (SCI) Laboratory, Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China
[2]Smart Computational Imaging Research Institute (SCIRI) of Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210019, China
[3]Jiangsu Key Laboratory of Visual Sensing & Intelligent Perception, Nanjing, Jiangsu Province 210094, China
[4]State Key Laboratory of Extreme Environment Optoelectronic Dynamic Measurement Technology and Instrument, Taiyuan, Shanxi Province, 030051, China
[5]Zircon Optoelectronics (Suzhou) Co., Ltd., Suzhou, Jiangsu Province, 215000, China
[*]chenqian@njust.edu.cn
[*]zuochao@njust.edu.cn
[†]Equal contribution

ABSTRACT

This document provides supplementary information for ”Transport-of-intensity diffraction tomography with nonuniform axial sampling (TIDT-NAS)”.

Contents

Supplementary Note 1. Derivation of the 3D partially coherent transfer function for transmission bright-field microscopy
Supplementary Note 2. Derivation of the discretely sampled spectrum in TIDT
Supplementary Note 3. Effect of axial scanning range on spectral sampling
Supplementary Note 4. Spectral aliasing and effective PTF support
Supplementary Note 5. Generation of the effective alias-free support for stack #2
Supplementary Note 6. Benchmark comparison between TIDT and TIDT-NAS

Supplementary Note 1. Derivation of the 3D partially coherent transfer function for transmission bright-field microscopy

The 3D imaging model for the phase tomographic microscopy starts from the theory of scalar diffraction tomography proposed by Wolf at 1969[1]. A physical quantity in the theory to describe the spatial distribution of the sample is scattering potential $f(\mathbf{x})$, which is represented by sample's and medium refractive index $n(\mathbf{x})$ and n_{m}, as:

$$f(\mathbf{x}) = k_0^2 \left[n^2(\mathbf{x}) - n_{\mathrm{m}}^2 \right] \tag{S1.1}$$

where $\mathbf{x} = (x, y, z)$ denotes the spatial coordinates, and $k_0 = 2\pi/\lambda$ is the wavenumber in free space. The $n(\mathbf{x})$ is a complex function whose real and imaginary parts respectively represent the refractive and absorptive characteristics. Through the Helmholtz equation satisfying the inhomogeneous medium, the relationship between the $f(\mathbf{x})$ and the 3D light field distribution $U(\mathbf{x})$ can be represented as:

$$\left[\nabla^2 + k_{\mathrm{m}}^2 \right] U(\mathbf{x}) = -f(\mathbf{x}) U(\mathbf{x}) \tag{S1.2}$$

where $k_{\mathrm{m}} = k_0 n_{\mathrm{m}}$ is the wavenumber in the medium. For the 3D light field $U(\mathbf{x})$, its composition can be separated into the incident light and the resultant diffracted light, represented as:

$$U(\mathbf{x}) = U_{\mathrm{in}}(\mathbf{x}) + U_{\mathrm{s}}(\mathbf{x}) \tag{S1.3}$$

The incident light field $U_{\mathrm{in}}(\mathbf{x})$ satisfies the Helmholtz equation in the homogeneous medium:

$$\left[\nabla^2 + k_{\mathrm{m}}^2 \right] U_{\mathrm{in}}(\mathbf{x}) = 0 \tag{S1.4}$$

Substituting Eqs. S1.3 and S1.4 into Eq. S1.2, it can be obtained that the diffracted light field satisfies the following relationship:

$$\left[\nabla^2 + k_{\mathrm{m}}^2 \right] U_{\mathrm{s}}(\mathbf{x}) = -f(\mathbf{x}) U(\mathbf{x}) \tag{S1.5}$$

Through the Green's theorem, the solution of the differential equation in Eq. S1.5 can be represented as:

$$U_{\mathrm{s}}(\mathbf{x}) = \int G(\mathbf{x} - \mathbf{x}_1) f(\mathbf{x}_1) U(\mathbf{x}_1) \, d\mathbf{x}_1 = \left[f(\mathbf{x}) U(\mathbf{x}) \right] \otimes G(\mathbf{x}) \tag{S1.6}$$

where $\otimes$ is the 3D convolution. $G(\mathbf{x})$ is the Green's function of the Helmholtz equation in the homogeneous medium, given by a spherical wave from a point with the spatial coordinate of $\mathbf{x}$:

$$G(\mathbf{x}) = \frac{e^{ik_{\mathrm{m}}|\mathbf{x}|}}{4\pi|\mathbf{x}|} \tag{S1.7}$$

Through dividing the $f(\mathbf{x})U(\mathbf{x})$, the Eq. S1.6 can be further expanded as:

$$U_{\mathrm{s}}(\mathbf{x}) = \int G(\mathbf{x} - \mathbf{x}_1) f(\mathbf{x}_1) U_{\mathrm{in}}(\mathbf{x}_1) \, d\mathbf{x}_1 + \int G(\mathbf{x} - \mathbf{x}_1) f(\mathbf{x}_1) U_{\mathrm{s}}(\mathbf{x}_1) \, d\mathbf{x}_1 \tag{S1.8}$$

If neglecting the multiple scattering in the interaction between the light and the material, the first-order scattering field can be represented as:

$$U_{s1}(\mathbf{x}) = \int G(\mathbf{x}-\mathbf{x}_1) f(\mathbf{x}_1) U_{in}(\mathbf{x}_1)\, d\mathbf{x}_1 \tag{S1.9}$$

Under the first-order Born approximation, the scattered field can be approximately equal to the first-order scattered field: $U_s(\mathbf{x}) \approx U_{s1}(\mathbf{x})$, hence, the total field can be represented as:

$$\begin{aligned} U(\mathbf{x}) &= U_{in}(\mathbf{x}) + U_s(\mathbf{x}) \\ &= U_{in}(\mathbf{x}) + \int G(\mathbf{x}-\mathbf{x}_1) f(\mathbf{x}_1) U_{in}(\mathbf{x}_1)\, d\mathbf{x}_1 \\ &= \int G(\mathbf{x}-\mathbf{x}_1) t(\mathbf{x}_1) U_{in}(\mathbf{x}_1)\, d\mathbf{x}_1 \end{aligned} \tag{S1.10}$$

where $t(\mathbf{x}) = 1 + f(\mathbf{x})$ can be seen as the 3D complex transmission function of the sample. Applying the Fourier transform to Eq. S1.10 yields:

$$U(\mathbf{x}) = \iint U_{in}(\mathbf{u})\, P(\mathbf{u}_1 - \mathbf{u})\, T(\mathbf{u}_1)\, e^{i\mathbf{u}_1\mathbf{x}}\, d\mathbf{u}\, d\mathbf{u}_1 \tag{S1.11}$$

where $\mathbf{u} = (u_x, u_y, u_z)$ is the spatial frequency coordinate. $U_{in}(\mathbf{u})$ is the set of 3D spatial-frequency coordinates corresponding to the incident angle of the light, and the $P(\mathbf{u})$ is the 3D CTF of the system, which are presented as:

$$\begin{aligned} P(\mathbf{u}) &= P(\mathbf{u}_T)\,\delta\left(u_z - \sqrt{\left(\frac{n_m}{\lambda}\right)^2 - |\mathbf{u}_T|^2}\right) \\ U_{in}(\mathbf{u}) &= U_{in}(\mathbf{u}_T)\,\delta\left(u_z - \sqrt{\left(\frac{n_m}{\lambda}\right)^2 - |\mathbf{u}_T|^2}\right) \end{aligned} \tag{S1.12}$$

where $P(\mathbf{u}_T)$ is the objective pupil function, and $U_{in}(\mathbf{u}_T)$ is the complex amplitude at the pupil plane of the condenser (equal to $\delta(\mathbf{u})$ under the coherent condition). Let us consider the partially coherent condition, where the incident light is generated by a quasi-monochromatic, incoherent extended source, represented as $U_{in}(\mathbf{u}_T, t_0)$. The phase difference between any two source points is not constant but varies randomly over time. Therefore, the time coordinate t_0 cannot be omitted here, and the light field is delta-correlated, satisfying:

$$\left\langle U_{in}(\mathbf{u}_{T,1}, t_0)\, U_{in}^*(\mathbf{u}_{T,2}, t_0)\right\rangle = \left|U_{in}(\mathbf{u}_{T,1})\right|^2 \delta(\mathbf{u}_{T,1} - \mathbf{u}_{T,2}) \tag{S1.13}$$

where $< \cdot >$ is the time averaging. The Eq. S1.11 can be rewritten under the partially coherent condition, as:

$$U(\mathbf{x}, t_0) = \iint U_{in}(\mathbf{u}, t_0)\, P(\mathbf{u}_1 - \mathbf{u})\, T(\mathbf{u}_1)\, e^{i\mathbf{u}_1\mathbf{x}}\, d\mathbf{u}_1\, d\mathbf{u} \tag{S1.14}$$

And the intensity measurement can be given by taking the time average of the resulting field intensity:

$$\begin{aligned} I(\mathbf{x}) &= \langle U(\mathbf{x}, t_0)\, U^*(\mathbf{x}, t_0)\rangle \\ &= \iiint P(\mathbf{u}+\mathbf{u}_1)\, P^*(\mathbf{u}+\mathbf{u}_2)\, e^{i2\pi\mathbf{x}(\mathbf{u}_1-\mathbf{u}_2)}\, d\mathbf{u}_1\, d\mathbf{u}_2\, d\mathbf{u} \end{aligned} \tag{S1.15}$$

where $S(\mathbf{u})$ is the intensity distribution of the light source at the condenser pupil plane, represented as:

$$S(\mathbf{u}) = \left|U_{\text{in}}(\mathbf{u}_{\text{T}})\right|^2 \delta\left(u_{\text{z}} - \sqrt{\left(\frac{n_{\text{m}}}{\lambda}\right)^2 - |\mathbf{u}_{\text{T}}|^2}\right) \tag{S1.16}$$

Separating the contributions of the specimen and the optical system leads to the concept of the transmission cross-coefficient (TCC)[2,3]:

$$TCC(\mathbf{u}_1, \mathbf{u}_2) = \int S(\mathbf{u})\, P(\mathbf{u}+\mathbf{u}_1)\, P^*(\mathbf{u}+\mathbf{u}_2)\, d\mathbf{u} \tag{S1.17}$$

and the partially coherent intensity stack can be written as:

$$I(\mathbf{x}) = \iint TCC(\mathbf{u}_1, \mathbf{u}_2)\, e^{i2\pi \text{x}(\mathbf{u}_1 - \mathbf{u}_2)}\, d\mathbf{u}_1\, d\mathbf{u}_2 \tag{S1.18}$$

Through separating the real and imaginary parts from the 3D complex transmission function, we can obtain:

$$t(\mathbf{x}) = 1 + f(\mathbf{x}) = 1 + k_0^2\left[n^2(\mathbf{x}) - n_{\text{m}}^2\right] = 1 + \phi(\mathbf{x}) + i\,a(\mathbf{x}) \tag{S1.19}$$

The real and imaginary parts $\phi(\mathbf{x})$ and $a(\mathbf{x})$ respectively correspond to the phase and absorption component of the object function. Under the first-order Born approximation, the mutual spectrum in Eq. S1.15 can be represented as:

$$T(\mathbf{u}_1)T^*(\mathbf{u}_2) = \delta(\mathbf{u}_1)\delta(\mathbf{u}_2) + \delta(\mathbf{u}_2)\left[\Phi(\mathbf{u}_1) + iA(\mathbf{u}_1)\right] + \delta(\mathbf{u}_1)\left[\Phi^*(\mathbf{u}_2) - iA^*(\mathbf{u}_2)\right] \tag{S1.20}$$

Substituting Eq. S1.20 into Eq. S1.15, the intensity can be reformulated as:

$$I(\mathbf{x}) = TCC(0,0) + 2Re\left\{\int TCC(\mathbf{u},0)\left[\Phi(\mathbf{u}) + iA(\mathbf{u})\right] e^{i2\pi\mathbf{x}\mathbf{u}}\, d\mathbf{u}\right\} \tag{S1.21}$$

where $TCC(\mathbf{u},0)$ is the linear part of the transmission cross-coefficient, known as the 3D weak object transfer function (WOTF):

$$\begin{aligned} WOTF(\mathbf{u}) &= TCC(\mathbf{u},0) = \int S(\mathbf{u}_1)\, P(\mathbf{u}_1+\mathbf{u})\, P^*(\mathbf{u}_1)\, d\mathbf{u}_1 \\ &= P(-\mathbf{u}) \otimes \left[S(\mathbf{u})\, P^*(\mathbf{u})\right] \end{aligned} \tag{S1.22}$$

Through separating the contribution of absorption and phase in intensity, Eq. S1.21 can be formulated as:

$$I(\mathbf{u}) = I_0\delta(\mathbf{u}) + (\mathbf{u})H_{\text{a}}(\mathbf{u})A(\mathbf{u}) + iH_\phi(\mathbf{u})\Phi(\mathbf{u}) \tag{S1.23}$$

Since the intensity is a real function, the real and even component of $WOTF(\mathbf{u})$ corresponds to the partially coherent 3D absorption transfer function (denoted as $H_{\text{a}}(\mathbf{u})$), and the imaginary and odd component of $WOTF(\mathbf{u})$ corresponds to the 3D phase transfer function (denoted as $H_\phi(\mathbf{u})$) of the object[4–6]:

$$\begin{aligned} H_{\text{a}}(\mathbf{u}) &= WOTF(\mathbf{u}) + WOTF^*(-\mathbf{u}) \\ &= P(-\mathbf{u}) \otimes \left[S(\mathbf{u})P^*(\mathbf{u})\right] + P^*(\mathbf{u}) \otimes \left[S(-\mathbf{u})P(-\mathbf{u})\right] \\ H_\phi(\mathbf{u}) &= WOTF(\mathbf{u}) - WOTF^*(-\mathbf{u}) \\ &= P(-\mathbf{u}) \otimes \left[S(\mathbf{u})P^*(\mathbf{u})\right] - P^*(\mathbf{u}) \otimes \left[S(-\mathbf{u})P(-\mathbf{u})\right] \end{aligned} \tag{S1.24}$$

Supplementary Note 2. Derivation of the discretely sampled spectrum in TIDT

We begin with the expression for the discretely sampled intensity stack $I_s(\mathbf{x}_T, z)$ in the spatial domain, given by:

$$I_s(\mathbf{x}_T, z) = I(\mathbf{x}_T, z)\ \mathrm{Comb}_{\Delta z}(z)\ \mathrm{rect}\left(\frac{z}{R}\right) \tag{S2.1}$$

where $\mathbf{x}_T = (x, y)$ denotes the transverse coordinates, Δz is the z-axial sampling interval, and R is the total axial scanning range. $I(\mathbf{x}_T, z)$ is the continuous signal of the intensity of the light field. The term $\mathrm{rect}(z/R)$ defines the finite scanning window $|z| < R/2$, while $\mathrm{Comb}_{\Delta z}(z)$ is the Dirac comb function representing discrete axial sampling:

$$\mathrm{Comb}_{\Delta z}(z) = \sum_{n=-\infty}^{+\infty} \delta(z - n\Delta z) \tag{S2.2}$$

Thus, the sampled intensity stack is obtained by multiplying the continuous intensity distribution by a periodic impulse train in z-axis and simultaneously truncating it within a finite axial range. To analyze how this discrete sampling affects the spectrum, we take the Fourier transform of Eq. S2.1 with respect to the axial coordinate z. Since multiplication in the spatial domain corresponds to convolution in the frequency domain, we write:

$$\begin{aligned} I_s(\mathbf{x}_T, u_z) &= \mathscr{F}_z\left\{ I(\mathbf{x}_T, z)\ \mathrm{Comb}_{\Delta z}(z)\ \mathrm{rect}\left(\frac{z}{R}\right) \right\} \\ &= I(\mathbf{x}_T, u_z) \otimes \mathscr{F}_z\{\mathrm{Comb}_{\Delta z}(z)\} \otimes \mathscr{F}_z\left\{ \mathrm{rect}\left(\frac{z}{R}\right) \right\} \end{aligned} \tag{S2.3}$$

The middle term in Eq. S2.3, i.e. the Fourier transform of the Dirac comb, can be expanded as:

$$\begin{aligned} \mathscr{F}_z\{\mathrm{Comb}_{\Delta z}(z)\} &= \sum_{n=-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(z - n\Delta z)\, e^{-i2\pi u_z z}\, dz \\ &= \sum_{n=-\infty}^{\infty} e^{-i2\pi u_z n\Delta z} \end{aligned} \tag{S2.4}$$

Based on the Poisson summation formula, Eq. S2.4 can be rewritten as:

$$\sum_{n=-\infty}^{\infty} e^{-i2\pi u_z n\Delta z} = \frac{1}{\Delta z} \sum_{n=-\infty}^{\infty} \delta\left(u_z - \frac{n}{\Delta z}\right) \tag{S2.5}$$

Since Eq. S2.5 describes an even function with respect to summation over all integer indices, the series can be equivalently expressed as:

$$\sum_{n=-\infty}^{\infty} \delta\left(-\Delta z\, u_z - n\right) = \sum_{n=-\infty}^{\infty} \delta\left(\Delta z\, u_z - n\right) \tag{S2.6}$$

Based on the scaling property of the Dirac delta function, we can obtain:

$$\delta\left(\Delta z\, u_z - n\right) = \frac{1}{\Delta z}\, \delta\left(u_z - \frac{n}{\Delta z}\right) \tag{S2.7}$$

Substituting Eqs. S2.5 - S2.7 into Eq. S2.4, the Fourier transform of the Dirac comb function can be represented as:

$$\mathscr{F}_z\{\mathrm{Comb}_{\Delta z}(z)\} = \frac{1}{\Delta z}\sum_{n=-\infty}^{\infty}\delta\left(u_z - \frac{n}{\Delta z}\right) \tag{S2.8}$$

The third term in Eq. S2.3, i.e. the Fourier transform of the finite window function, can be represented as:

$$\begin{aligned}\mathscr{F}_z\left\{\mathrm{rect}\left(\frac{z}{R}\right)\right\} &= \int_{-\infty}^{\infty}\mathrm{rect}\left(\frac{z}{R}\right)e^{-i2\pi u_z z}\,dz \\ &= \int_{-R/2}^{R/2} e^{-i2\pi u_z z}\,dz\end{aligned} \tag{S2.9}$$

By evaluating the definite integral in Eq. S2.9, this term is finally expressed as a sinc function:

$$\begin{aligned}\mathscr{F}_z\left\{\mathrm{rect}\left(\frac{z}{R}\right)\right\} &= \left.\frac{e^{-i2\pi u_z z}}{-i2\pi u_z}\right|_{-R/2}^{R/2} \\ &= \frac{e^{-i\pi u_z R} - e^{i\pi u_z R}}{-i2\pi u_z} \\ &= \frac{\sin(\pi u_z R)}{\pi u_z} \\ &= R\cdot\mathrm{sinc}\left(R u_z\right)\end{aligned} \tag{S2.10}$$

Substituting Eqs. S2.8 and S2.10 into Eq. S2.3, we finally obtain an explicit expression for the spectrum of the axially sampled intensity stack. It shows that the continuous axial spectrum is replicated at intervals of $1/\Delta z$ along u_z, due to the Dirac comb sampling, and each replicated copy is further convoluted by a kernel with the distribution of a sinc function:

$$I_s(\mathbf{x}_T, u_z) = \frac{R}{\Delta z}\left[\sum_{n=-\infty}^{\infty} I\left(\mathbf{x}_T, u_z - \frac{n}{\Delta z}\right)\right] \otimes \mathrm{sinc}\left(R u_z\right) \tag{S2.11}$$

Finally, we take the 2D Fourier transform with respect to the transverse coordinate $\mathbf{x}_T$. Since the transverse Fourier transform is independent of the axial convolution and shifting operations, we obtain:

$$I_s(\mathbf{u}_T, u_z) = \frac{R}{\Delta z}\left[\sum_{n=-\infty}^{\infty} I\left(\mathbf{u}_T, u_z - \frac{n}{\Delta z}\right)\right] \otimes \mathrm{sinc}\left(R u_z\right) \tag{S2.12}$$

Supplementary Note 3. Effect of axial scanning range on spectral sampling

Equation S2.12 shows that a finite axial scanning range introduces a convolution of the intensity spectrum with a $\mathrm{sinc}(Ru_z)$ kernel, whose main-lobe width is approximately $2/R$. Therefore, reducing the axial scanning range broadens the convolution kernel in the frequency domain and decreases the effective axial spectral resolution. To evaluate this effect, we compare three through-focus intensity stacks acquired with the same axial step size of 0.32 μm but different scanning ranges of $R = 40$ μm, $R = 20$ μm, and $R = 10$ μm. A water-immersion objective lens (40×, 1.1 NA) is used in this experiment. The corresponding intensity spectra and PTFs are shown in Fig. S1a, and the raw intensity stacks in the x-z plane are shown in Fig. S1c.

We define the discretely sampled spectrum of the first intensity stack, acquired over the largest scanning range, as

$$I_{\mathrm{s},1}\left[\mathbf{u}_{\mathrm{T}}, k\right], \tag{S3.1}$$

where k is the discrete axial-frequency index. This index can be mapped to the corresponding axial spatial-frequency coordinate as

$$\begin{aligned} k &= \left\{-\frac{N}{2}, -\frac{N}{2}+1, \ldots, 0, \ldots, \frac{N}{2}-1, \frac{N}{2}\right\}, \\ \frac{k}{N\Delta z} &= \left\{-\frac{N}{2N\Delta z}, -\frac{N-2}{2N\Delta z}, \ldots, 0, \ldots, \frac{N-2}{2N\Delta z}, \frac{N}{2N\Delta z}\right\}. \end{aligned} \tag{S3.2}$$

Here, N denotes the total number of axial layers in the first stack. Since the axial step size is sufficiently small to satisfy the Nyquist criterion, spectral duplication caused by axial undersampling is negligible in this analysis and is not considered in Eq. S3.1.

When the scanning range is reduced while keeping the same axial step size, the axial spectral interval becomes larger. Equivalently, the spectrum is convolved with a broader sinc kernel and then sampled on a coarser frequency grid. For the second and third stacks, corresponding to scanning ranges of 20 μm and 10 μm, respectively, their spectra can be expressed as weighted sums of the spectral samples in the first stack:

$$I_{\mathrm{s},2}\left[\mathbf{u}_{\mathrm{T}}, k\right] = \frac{1}{2}\sum_{n=0}^{N-1}\left(I_{\mathrm{s},1}\left[\mathbf{u}_{\mathrm{T}}, n+1\right]\mathrm{sinc}\left[\frac{2k-n}{2}\right]\right), \quad k = -\frac{N}{4}, -\frac{N}{4}+1, \ldots, 0, \ldots, \frac{N}{4}-1, \frac{N}{4}, \tag{S3.3}$$

$$I_{\mathrm{s},3}\left[\mathbf{u}_{\mathrm{T}}, k\right] = \frac{1}{4}\sum_{n=0}^{N-1}\left(I_{\mathrm{s},1}\left[\mathbf{u}_{\mathrm{T}}, n+1\right]\mathrm{sinc}\left[\frac{4k-n}{4}\right]\right), \quad k = -\frac{N}{8}, -\frac{N}{8}+1, \ldots, 0, \ldots, \frac{N}{8}-1, \frac{N}{8}. \tag{S3.4}$$

Eqs. S3.3 and S3.4 indicate that each spectral pixel in the reduced-range acquisition is not a direct sample of the original spectrum, but a weighted average of neighboring spectral pixels. This effect is schematically illustrated in Fig. S1b. As the scanning range decreases, the corresponding sinc kernel becomes wider, causing each spectral pixel to integrate information from a broader axial-frequency neighborhood. In practice, the 20 μm and 10 μm scanning ranges mainly integrate approximately 3 and

7 original spectral pixels, respectively. This spectral pixel binning reduces the effective axial spectral resolution and preferentially degrades low-frequency volumetric information, which is essential for recovering global RI contrast.

The RI reconstructions in Figs. S1d and S1e confirm this behavior. When the scanning range is reduced from 40 μm to 20 μm and 10 μm, the reconstructed *x*-*z* slices show progressively weakened low-frequency volumetric contrast, while the *x*-*y* slices exhibit reduced global RI fidelity. High-frequency subcellular details can still be partially preserved under a short scanning range because these features are primarily encoded near the in-focus region. However, the slowly varying low-frequency information originates from large-defocus diffraction contrast and therefore requires a sufficiently extended axial range to be recorded. Once the scanning range is truncated, this part of the wavefront information is lost in the spatial domain and appears as spectral smoothing or binning in the frequency domain, leading to degraded low-frequency RI reconstruction.

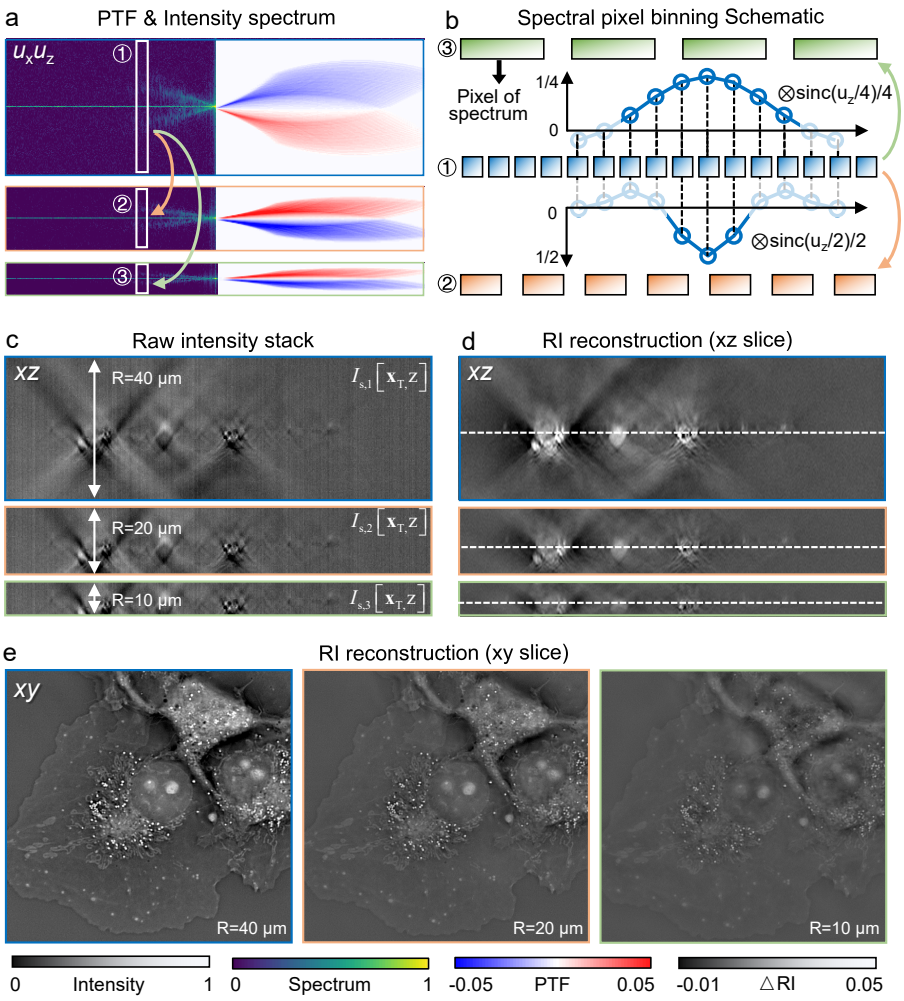


Figure S1. Effect of axial scanning range on discretely sampled spectra and RI reconstruction. a. Measured intensity spectra and corresponding PTFs in the u_x-u_z plane for three axial scanning ranges, denoted as cases ①–③. **b.** Schematic illustration of spectral pixel binning along the u_z axis induced by finite axial-range truncation. A shorter scanning range broadens the corresponding sinc kernel and results in stronger spectral averaging, thereby reducing the effective axial spectral resolution. **c.** Raw through-focus intensity stacks shown in the x-z plane for axial scanning ranges of $R = 40\ \mu$m, $R = 20\ \mu$m, and $R = 10\ \mu$m. **d.** RI reconstructions in the x-z plane obtained from the datasets in **c**. The dashed lines indicate the axial positions of the x-y slices shown in **e**. **e.** Corresponding RI reconstructions in the x-y plane.

Supplementary Note 4. Spectral aliasing and effective PTF support

To further examine the influence of the axial sampling step Δz on spectral acquisition in TIDT, we compare intensity stacks acquired with step sizes of 0.32 μm, 0.64 μm, and 1.28 μm over the same axial scanning range of 40 μm. According to Eq. S2.12 and Eq. S3.1, discrete axial sampling causes the continuous spectrum to be periodically replicated along the axial frequency axis. For the well-sampled reference stack, the sampled spectrum can be written as

$$I'_{\mathrm{s},1}\left[\mathbf{u}_{\mathrm{T}},k\right]=\sum_{n=-\infty}^{\infty}I_{\mathrm{s},1}\left[\mathbf{u}_{\mathrm{T}},k+nN\right],\quad k=-\frac{N}{2},\,-\frac{N}{2}+1,\,\ldots,\,0,\,\ldots,\,\frac{N}{2}-1,\,\frac{N}{2}. \tag{S4.1}$$

Here, $I_{\mathrm{s},1}\left[\mathbf{u}_{\mathrm{T}},k\right]$ denotes the discrete spectrum along the axial frequency index k, and the period N corresponds to the axial sampling rate $1/\Delta z$. To avoid aliasing, the sampling rate must be larger than twice the maximum axial spatial frequency supported by the 3D PTF. This maximum axial frequency can be estimated from the axial cutoff of the transmissive bright-field microscope[7]:

$$u_z^{\mathrm{cutoff}}=\frac{n_{\mathrm{m}}-\sqrt{n_{\mathrm{m}}^2-\mathrm{NA}_{\mathrm{obj}}^2}}{\lambda_0}, \tag{S4.2}$$

where $\mathrm{NA}_{\mathrm{obj}}$ is the numerical aperture of the objective lens, n_{m} is the refractive index of the surrounding medium, and λ_0 is the wavelength in vacuum. For $\Delta z=0.32$ μm with a 40×/1.1 NA objective lens, the condition $1/\Delta z>2u_z^{\mathrm{cutoff}}$ is satisfied. Therefore, the spectral replicas are separated along the u_z direction, and the complete PTF-supported spectrum is recorded without aliasing, as shown in Fig. S2a1. The corresponding duplicated spectral supports and PTFs in Fig. S2b1 further confirm that the axial sampling rate is sufficient to cover the full PTF support. In general, this criterion provides a practical reference for selecting the dense axial step size in conventional TIDT and in stack #1 of TIDT-NAS.

When the axial step size is doubled or quadrupled, the axial sampling rate decreases accordingly, causing the spectral replicas to move closer to each other. The corresponding sampled spectra can be formulated as

$$I'_{\mathrm{s},2}\left[\mathbf{u}_{\mathrm{T}},k\right]=\sum_{n=-\infty}^{\infty}I_{\mathrm{s},1}\left[\mathbf{u}_{\mathrm{T}},k+n\frac{N}{2}\right],\quad k=-\frac{N}{4},\,-\frac{N}{4}+1,\,\ldots,\,0,\,\ldots,\,\frac{N}{4}-1,\,\frac{N}{4}, \tag{S4.3}$$

and

$$I'_{\mathrm{s},3}\left[\mathbf{u}_{\mathrm{T}},k\right]=\sum_{n=-\infty}^{\infty}I_{\mathrm{s},1}\left[\mathbf{u}_{\mathrm{T}},k+n\frac{N}{4}\right],\quad k=-\frac{N}{8},\,-\frac{N}{8}+1,\,\ldots,\,0,\,\ldots,\,\frac{N}{8}-1,\,\frac{N}{8}. \tag{S4.4}$$

Equations S4.3 and S4.4 show that increasing the axial step size does not compress the underlying continuous spectrum itself; instead, it reduces the axial sampling bandwidth and the duplication interval in the discrete spectrum. As a result, the periodically replicated spectra become closer and begin to overlap within the sampled passband.

As shown in Fig. S2a2–a3, spectral aliasing appears when the axial step is doubled and becomes more pronounced when the step is quadrupled. The corresponding support analysis in Fig. S2b2–b3 shows that

the high-frequency components are first affected by aliasing, whereas part of the low-frequency support can remain non-overlapped and therefore usable for reconstruction. With further increase of the axial step size, the aliased regions progressively extend toward the mid-frequency components, leaving only a smaller alias-free low-frequency region.

This observation provides the basis for defining the effective PTF support used in TIDT-NAS. For the sparsely sampled stack #2, the periodically replicated PTF supports are accumulated along the u_z direction according to the sparse sampling period $1/\Delta z_s$. Regions covered by more than one replicated support correspond to aliased spectral components and are excluded, whereas the single-covered low-frequency region is retained as the effective alias-free support $\Omega_{\text{eff}}(\mathbf{u})$. Therefore, although axial undersampling introduces high-frequency aliasing, the remaining alias-free low-frequency spectrum of stack #2 can still be reliably fused with the high-frequency information provided by the densely sampled stack #1.

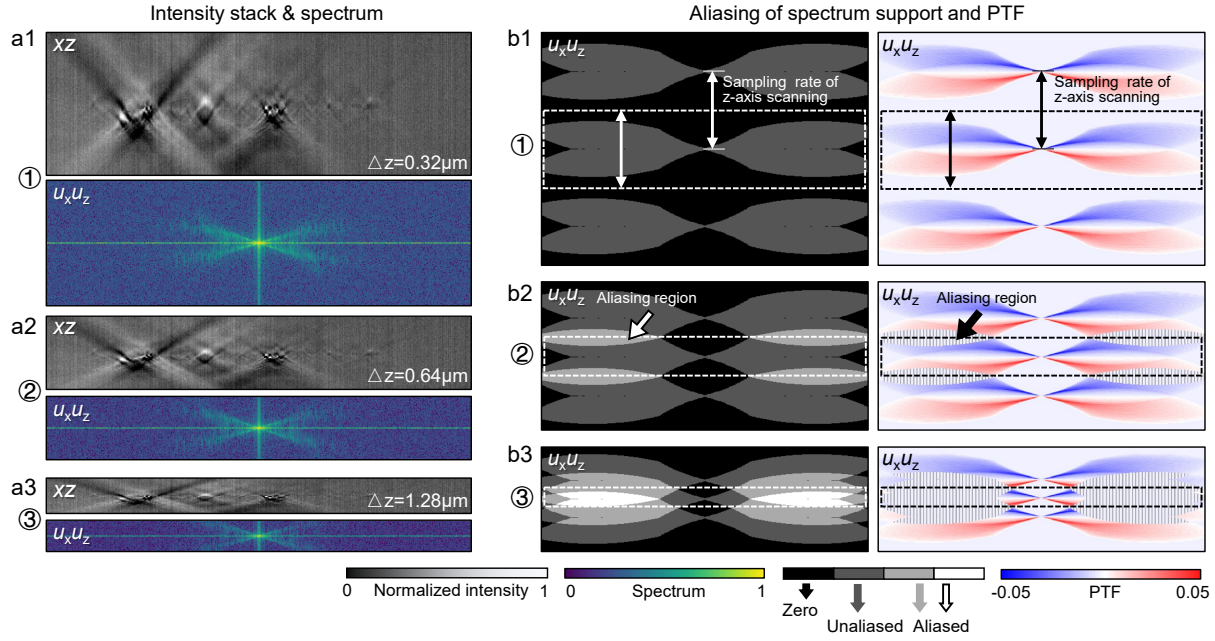


Figure S2. Schematic diagram of spectrum duplication and aliasing under the discrete z-axis sampling. a1-a3. The intensity stacks acquired under different scanning steps, 0.32μm, 0.64μm, and 1.28μm, as well as the corresponding intensity spectra. **b1-b3**. The supports of duplicated spectra corresponding to the spectra in (a) under different scanning steps, and the corresponding duplicated PTFs.

Supplementary Note 5. Generation of the effective alias-free support for stack #2

In TIDT-NAS, stack #2 is acquired over the full axial scanning range with an enlarged axial sampling step Δz_s. This sparse axial sampling preserves low-frequency volumetric information with fine axial spectral resolution, but also introduces axial aliasing due to the periodic replication of the 3D PTF support along the u_z direction. Therefore, before fusing the spectrum reconstructed from stack #2 with that from stack #1, it is necessary to identify the spectral region of stack #2 that remains free from aliasing. This region is referred to as the effective alias-free support, denoted as $\Omega_{\text{eff}}(\mathbf{u})$. The complete procedure for generating $\Omega_{\text{eff}}(\mathbf{u})$ is illustrated in Fig. S3.

We first compute the 3D PTF corresponding to stack #1 after zero-padding, denoted as $H_\phi^{\#1}(\mathbf{u})$ (Fig. S3a). Because stack #1 is densely sampled with the axial step Δz, its PTF support is treated as the reference support of the well-sampled system. The binary support of PTF #1 is then extracted as

$$\Omega_{\text{PTF}}(\mathbf{u}) = \left\{ \mathbf{u} \in \mathbb{R}^3 : \left| H_\phi^{\#1}(\mathbf{u}) \right| > \tau \right\}, \tag{S5.1}$$

where τ is a small threshold used to suppress numerical residuals outside the physical support. This binary support is shown in Fig. S3b. The dense axial sampling of stack #1 defines an axial sampling bandwidth of $1/\Delta z$, whereas the sparse axial sampling of stack #2 defines a narrower axial passband of $1/\Delta z_s$.

Under sparse axial sampling with step Δz_s, the PTF support is periodically replicated along the u_z direction with a period of $1/\Delta z_s$. Therefore, the shifted support regions can be expressed as

$$\Omega_{\text{PTF}}^{(n)}(\mathbf{u}) = \Omega_{\text{PTF}} \left(\mathbf{u}_\text{T}, u_z - \frac{n}{\Delta z_s} \right), \quad n \in \mathbb{Z}, \tag{S5.2}$$

where $\mathbf{u}_\text{T} = (u_x, u_y)$ is the transverse spatial-frequency coordinate. In the present imaging configuration, the dominant aliasing within the axial passband of stack #2 is captured by the three support regions corresponding to $n = -1, 0, 1$, as illustrated in Fig. S3c.

To determine whether a spectral location is contaminated by aliasing, the shifted support regions are accumulated:

$$C(\mathbf{u}) = \sum_{n=-1}^{1} \Omega_{\text{PTF}} \left(\mathbf{u}_\text{T}, u_z - \frac{n}{\Delta z_s} \right). \tag{S5.3}$$

The accumulated support map is shown in Fig. S3d. Regions with $C(\mathbf{u}) = 1$ are covered by only one replicated support and therefore remain alias-free. In contrast, regions with $C(\mathbf{u}) > 1$ are covered by multiple shifted supports and correspond to aliased spectral components. Regions with $C(\mathbf{u}) = 0$ lie outside the PTF support and do not contribute useful transferred information.

The candidate support of stack #2 is further restricted to its axial passband, determined by the sparse sampling rate:

$$|u_z| \leq \frac{1}{2\Delta z_s}. \tag{S5.4}$$

As shown in Fig. S3e, this operation crops the accumulated support within the spectral band accessible to stack #2. Finally, the effective alias-free support is obtained by retaining only the single-covered region within this passband:

$$\Omega_{\mathrm{eff}}(\mathbf{u}) = \begin{cases} 1, & \displaystyle\sum_{n=-1}^{1} \Omega_{\mathrm{PTF}}\left(\mathbf{u}_{\mathrm{T}}, u_z - \frac{n}{\Delta z_{\mathrm{s}}}\right) = 1, \\ 0, & \displaystyle\sum_{n=-1}^{1} \Omega_{\mathrm{PTF}}\left(\mathbf{u}_{\mathrm{T}}, u_z - \frac{n}{\Delta z_{\mathrm{s}}}\right) \neq 1. \end{cases} \tag{S5.5}$$

In practice, the above definition is applied within the axial passband of stack #2, while spectral regions outside the passband are assigned zero. The resulting $\Omega_{\mathrm{eff}}(\mathbf{u})$ is shown in Fig. S3f.

This effective support identifies the reliable low-frequency spectral components that can be extracted from the sparsely sampled stack #2. During the Fourier-domain fusion in TIDT-NAS, the spectrum reconstructed from stack #2 is retained only within $\Omega_{\mathrm{eff}}(\mathbf{u})$, whereas the remaining complementary spectral components are supplied by stack #1. In this way, the aliased components introduced by sparse axial sampling are rejected, while the useful low-frequency volumetric information acquired over the full scanning range is preserved for high-fidelity RI reconstruction.

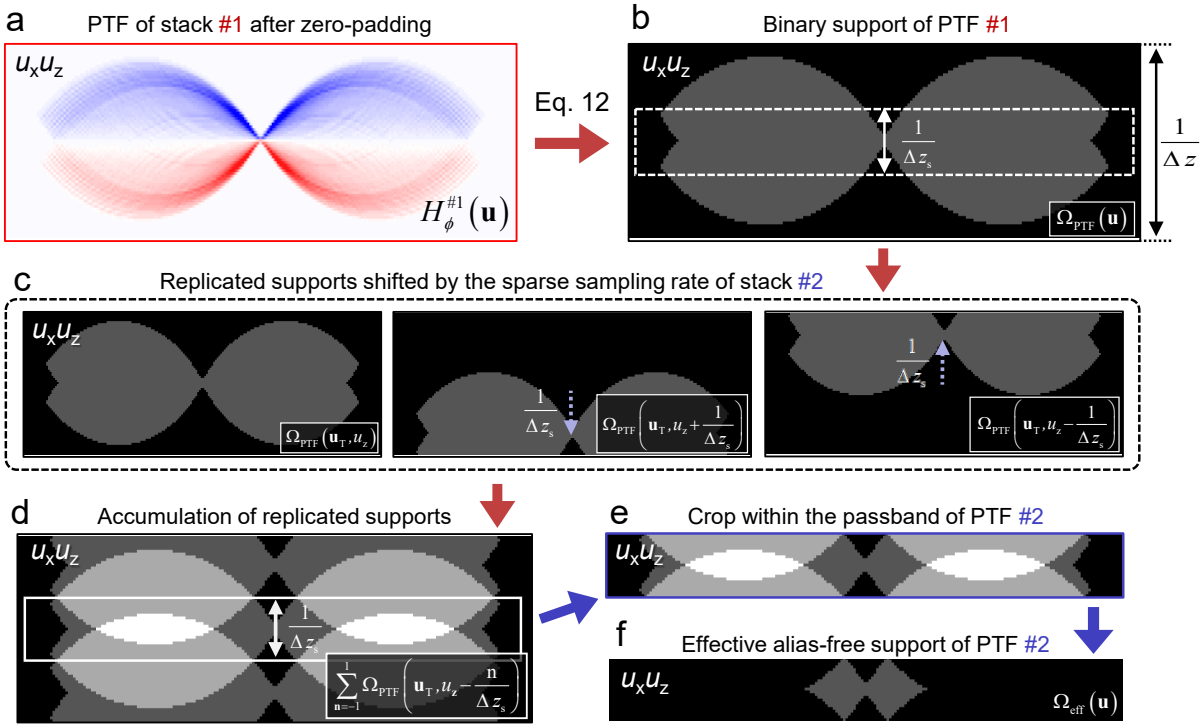


Figure S3. Generation of the effective alias-free support for stack #2. a. The 3D PTF of stack #1 after zero-padding, shown in the u_x-u_z plane. **b.** Binary support of PTF #1, denoted as $\Omega_{\mathrm{PTF}}(\mathbf{u})$. The axial frequency interval of the densely sampled stack is $1/\Delta z$, while the sparse sampling of stack #2 defines an axial passband of $1/\Delta z_s$. **c.** Periodically replicated PTF supports shifted along the u_z direction with period $1/\Delta z_s$, corresponding to the aliasing period introduced by sparse axial sampling. **d.** Accumulation of the replicated supports. Regions covered by more than one shifted support correspond to aliased spectral components, whereas single-covered regions remain alias-free. **e.** Candidate support within the axial passband of PTF #2. **f.** Effective alias-free support $\Omega_{\mathrm{eff}}(\mathbf{u})$ used to extract reliable low-frequency spectral components from stack #2 for the subsequent Fourier-domain fusion.

Supplementary Note 6. Benchmark comparison between TIDT and TIDT-NAS

To evaluate whether TIDT-NAS can achieve 3D RI reconstruction quality comparable to conventional TIDT while using a compressed dataset, we performed experiments on standard quantitative phase targets placed horizontally on the microscope stage. A 40×/0.75 NA objective lens was used for data acquisition. For conventional TIDT, the through-focus intensity stack was acquired over an axial range of ∼50 μm with a step size of 0.6 μm, resulting in 83 axial layers. For TIDT-NAS, two complementary stacks were acquired: stack #1 covered a reduced axial range of ∼10 μm with the dense step size of 0.6 μm, whereas stack #2 covered the full axial range of ∼50 μm with a sparse step size of 3 μm. In total, TIDT-NAS required only 27 axial images.

We first evaluated the reconstruction performance using a USAF-1951 target, as shown in Fig. S4a,b. The raw through-focus intensity stack is shown in the x-y plane and along two selected x-z sections. The corresponding RI reconstructions obtained by TIDT-NAS and conventional TIDT show highly consistent lateral and axial structures. In particular, the x-z sections along Line 1 and Line 2 demonstrate that TIDT-NAS preserves axial reconstruction fidelity comparable to that of conventional TIDT, despite the substantial reduction in the number of acquired images. The slight RI depression observed in the axial sections is attributed to the missing-cone effect of the scattering potential under limited illumination NA[8]. We further repeated the experiment using a Siemens star target to assess the generality of the reconstruction performance, as shown in Fig. S4c,d. The raw intensity stack and the reconstructed RI distributions again show close agreement between TIDT-NAS and conventional TIDT. The corresponding y-z sections confirm that TIDT-NAS maintains comparable axial tomographic fidelity for structures with different spatial orientations and frequency distributions.

The spectral characteristics of the two acquisition schemes are summarized in Fig. S4e–g. The PTFs of TIDT-NAS stack #1, TIDT-NAS stack #2, and conventional TIDT are shown in Fig. S4e, while the measured intensity spectra of the USAF-1951 and Siemens star targets are shown in Fig. S4f,g. The contours of the experimentally acquired object spectra agree well with the corresponding 3D PTF supports, confirming the accuracy of the system calibration and the reliability of the transfer-function-based reconstruction. In TIDT-NAS, stack #1 and stack #2 provide complementary spectral information: stack #1 preserves high-frequency structural components, whereas stack #2 retains full-range low-frequency volumetric information. Overall, the benchmark experiments on both targets verify that TIDT-NAS achieves 3D RI reconstructions comparable to conventional TIDT while reducing the number of acquired axial images from 83 to 27, providing objective validation of the proposed nonuniform axial sampling and Fourier-domain fusion strategy.

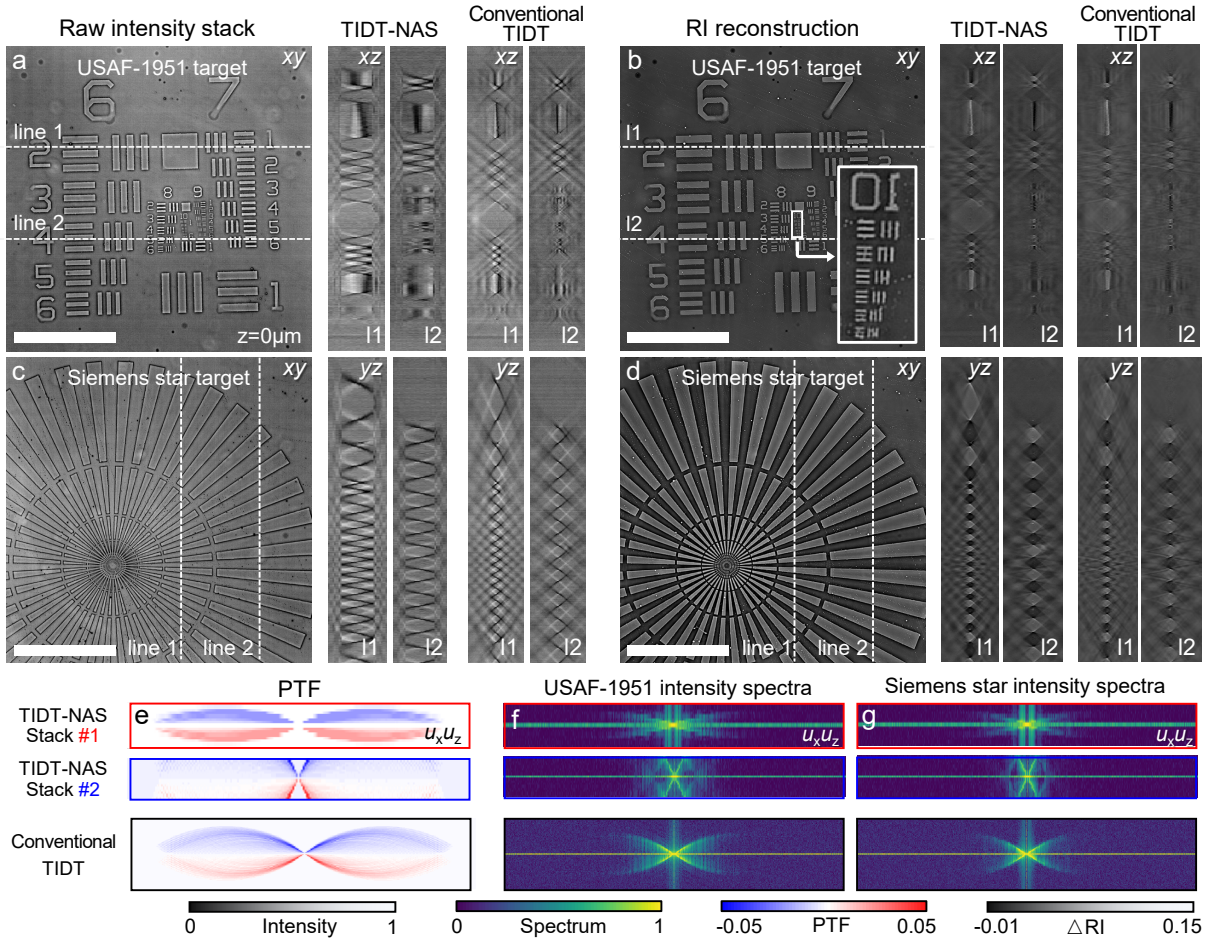
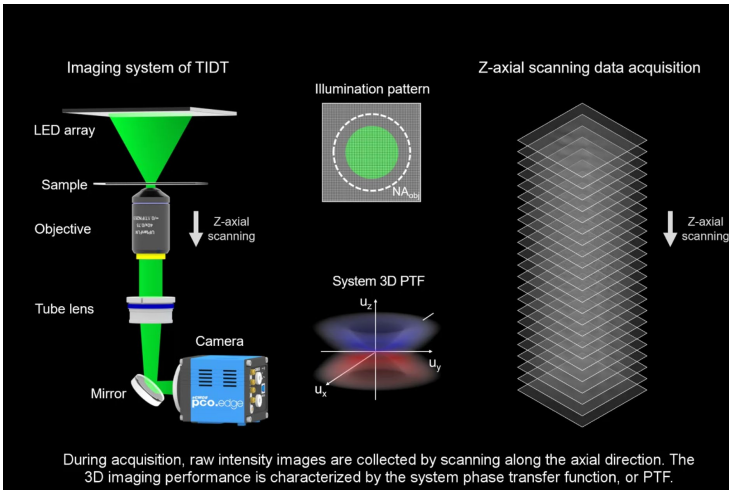
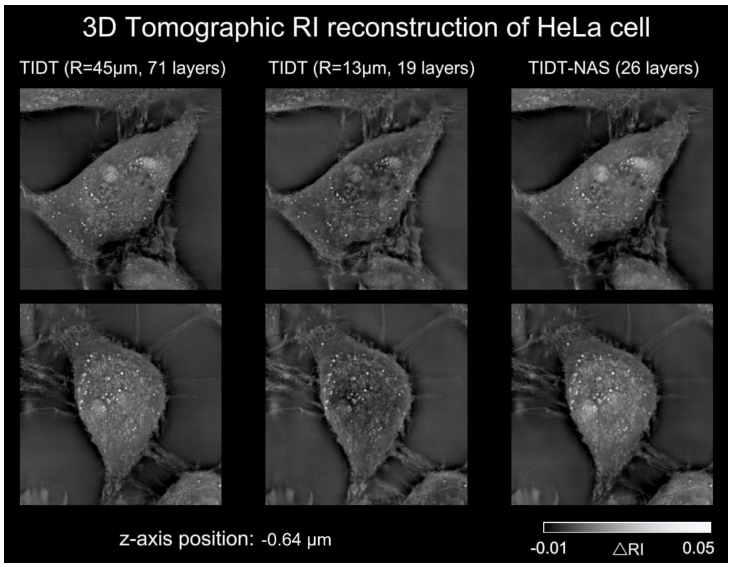


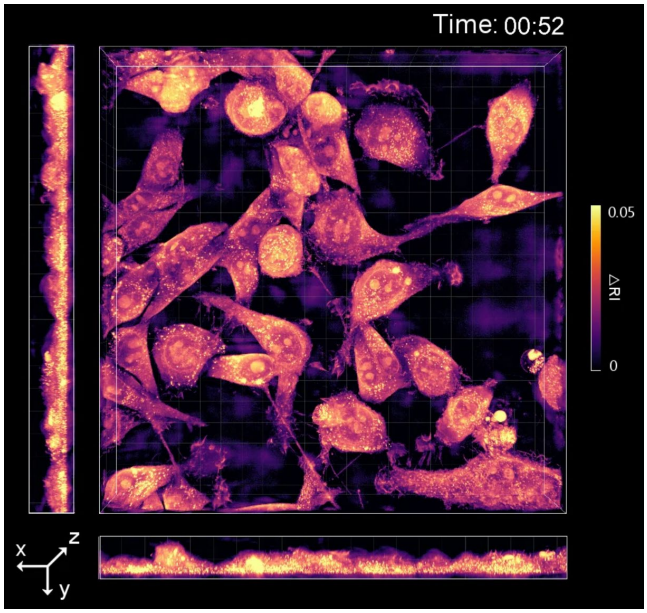
Figure S4. Validation of TIDT-NAS on standard phase targets. a,b. Raw intensity stack and RI reconstruction of the USAF-1951 target obtained by TIDT-NAS and conventional TIDT. The selected *x*-*z* sections show comparable axial reconstruction fidelity between the two methods. **c,d.** Raw intensity stack and RI reconstruction of the Siemens star target, with corresponding *y*-*z* sections comparing the tomographic reconstruction performance. **e.** 3D PTFs of TIDT-NAS stack #1, TIDT-NAS stack #2, and conventional TIDT. **f,g.** Intensity spectra of the USAF-1951 and Siemens star targets. The complementary spectra from the two TIDT-NAS stacks support RI reconstruction comparable to conventional TIDT while using fewer axial measurements. Scale bars: 20 μm.



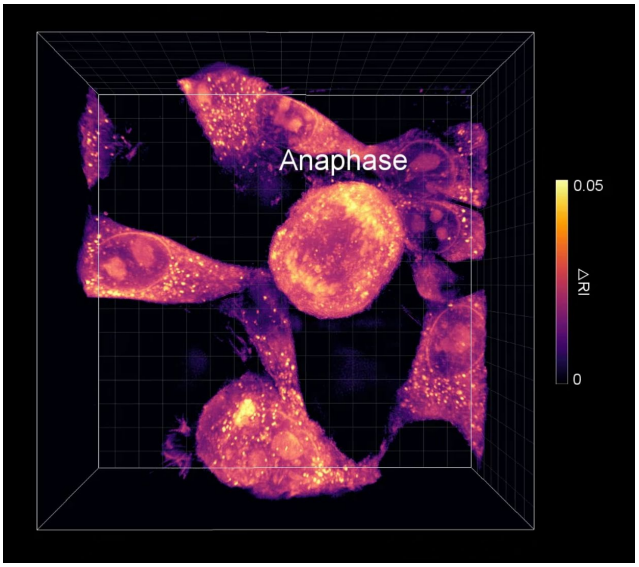
Visualization 1. Effects of axial scanning parameters on spectral sampling and TIDT-NAS reconstruction workflow



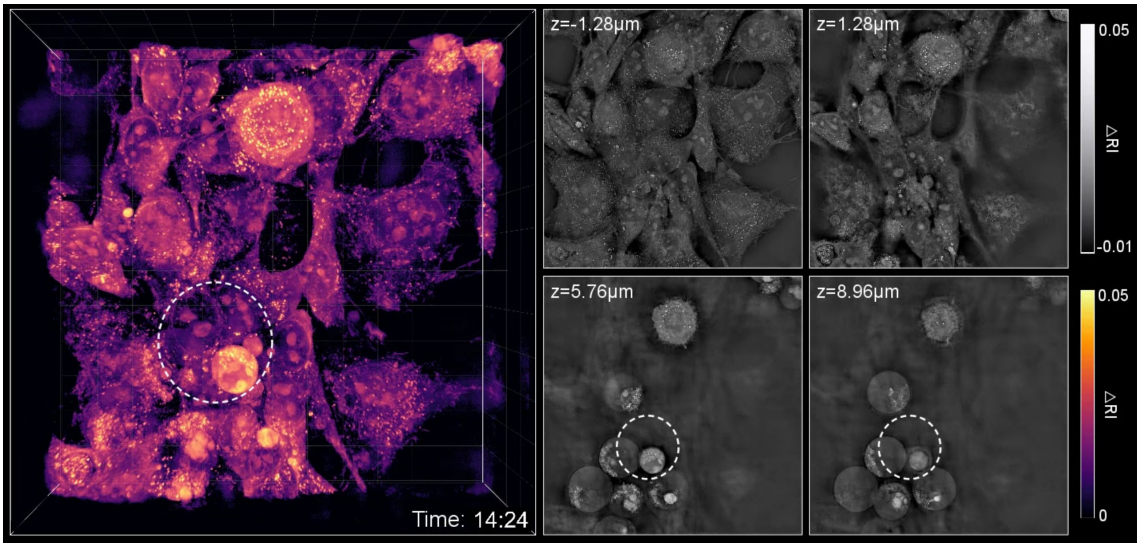
Visualization 2. Comparison of TIDT and TIDT-NAS for 3D RI tomography



Visualization 3. 16-hour RI volume rendering of in vitro HeLa cells



Visualization 4. RI volume rendering of cell mitosis process



Visualization 5. RI tomography and volume rendering for vacuolation of a single HeLa cell

References

1. Wolf, E. Three-dimensional structure determination of semi-transparent objects from holographic data. *Opt. communications* **1**, 153–156 (1969).
2. Sheppard, C. & Choudhury, A. Image formation in the scanning microscope. *Opt. Acta: Int. J. Opt.* **24**, 1051–1073 (1977).
3. Hopkins, H. H. On the diffraction theory of optical images. *Proc. Royal Soc. Lond. Ser. A. Math. Phys. Sci.* **217**, 408–432 (1953).
4. Sheppard, C. J. Three-dimensional phase imaging with the intensity transport equation. *Appl. optics* **41**, 5951–5955 (2002).
5. Soto, J. M., Rodrigo, J. A. & Alieva, T. Label-free quantitative 3d tomographic imaging for partially coherent light microscopy. *Opt. express* **25**, 15699–15712 (2017).
6. Zuo, C. *et al.* Transport of intensity equation: a tutorial. *Opt. Lasers Eng.* **135**, 106187 (2020).
7. Haeberlé, O., Belkebir, K., Giovaninni, H. & Sentenac, A. Tomographic diffractive microscopy: basics, techniques and perspectives. *J. Mod. Opt.* **57**, 686–699 (2010).
8. Lim, J. *et al.* Comparative study of iterative reconstruction algorithms for missing cone problems in optical diffraction tomography. *Opt. express* **23**, 16933–16948 (2015).