

Supplementary Information:

Bypassing the imaging thickness limit with a translationally-invariant nonlocal thin film

Bohan Zhang¹, Liu Li¹, Xinyuan Xue¹, Yibao Xu¹, Qitong Li², Hao Zhang¹, Liqun Sun¹, Liangcai Cao¹, Yuanmu Yang^{1*}

¹State Key Laboratory of Precision Measurement Technology and Instruments, Department of Precision Instrument, Tsinghua University, Beijing 100084, China

²Department of Electronic Engineering, Tsinghua University, Beijing 100084, China

*Email: ymyang@tsinghua.edu.cn

Supplementary Note 1. Elaborated principle of deterministic 2D imaging using the nonlocal thin film

The nonlocal thin film discussed in the main text is structured as a Bragg microcavity, which consists of a central half-wave cavity layer with refractive index n_L , sandwiched between two distributed Bragg reflectors (DBRs), each comprising M alternating quarter-wave layers of high-index n_H and low-index n_L dielectrics. This structure is denoted as $(LH)^M 2L (HL)^M$. For sufficiently large M , the multilayer stack approximates a Fabry–Pérot cavity that exhibits a narrowband, angle-dependent transmission response governed by interference of multiple internal reflections. For clarity, we restrict our analysis to the one-dimensional case. Its complex transmission as a function of the transverse spatial frequency f_x follows the Airy function as,

$$t(f_x) = \frac{(1 - \rho)e^{i\delta/2}}{1 - \rho e^{i\delta}}, \quad (S1)$$

where $\delta = 2\pi\sqrt{1 - \lambda^2 f_x^2 / n_L^2}$ is the phase delay accumulated over a round trip through the Fabry–Pérot cavity at wavelength λ , and $\rho \approx 1 - \frac{4n_L^{2M+1}}{n_H^{2M}}$ is the DBR's reflectivity. In the high-reflectivity limit, the Airy function in Eq. (S1) simplifies near the transmission resonance at normal incidence, up to a constant phase factor, to a complex Lorentzian form as,

$$t(f_x) = \frac{\tilde{f}_{x0}^2}{f_{x0}^2 + f_x^2}, \quad (S2)$$

where the characteristic complex spatial frequency is $\tilde{f}_{x0} = \frac{2e^{-i\frac{\pi}{4}} n_L^{M+1.5}}{\sqrt{\pi}\lambda n_H^M}$.

Via the inverse Fourier transform of $t(f_x)$, one can obtain the normalized Fabry–Pérot multi-beam interference kernel $h(x)$ of the Bragg cavity as,

$$h(x) = e^{-2\pi\tilde{f}_{x0}|x|}, \quad (S3)$$

which decays exponentially with the lateral distance between the input and output positions x . Eq. (S3) clearly indicates that the response of a Bragg microcavity is nonlocal. Physically, this kernel reflects the cumulative effect of multiple-beam interference, whereby light incident at one point undergoes a series of internal reflections that leak out at shifted positions, thereby spreading the output field over a finite lateral range.

The real part of $\tilde{f}_{x0}$ determines the spatial attenuation of the nonlocal response, with smaller values indicating broader lateral spreading. The imaginary part governs the co-phase condition across the film, where smaller values imply reduced optical path differences through the thin film. From Eq. (S2), the characteristic spatial frequency has a phase angle of $-\pi/4$. At the half-maximum amplitude position, the corresponding phase difference from the center is $\ln 2 = 0.693$ rad. This value is well within the Rayleigh quarter wavelength rule¹ for high-quality imaging limit of $\pi/2$ rad (corresponding to a $\lambda/4$ path difference), which supports the claim of constructive interference for light originating from within the lateral spreading length W .

For comparison, with the film's thickness of $12.2 \mu\text{m}$ and a lateral spreading W of $69.6 \mu\text{m}$ (for $M = 13$), the phase difference for free-space propagation over the same lateral distance would be approximately 99.0 rad. The film effectively reduces this substantial phase difference to 0.693 rad, enabling constructive interference of light from a continuous spatial range to a single output point without conventional lens-based propagation. It is also evident from Supplementary Fig. 1b and Eq. (S3) that the phase difference for a given lateral distance decreases as the cavity Q-factor increases, meaning a higher Q-factor further improves the co-phase condition.

To elucidate the interference mechanism, Supplementary Figure 2 illustrates the simulated electric field intensity for a one-dimensional system under cylindrical wave incidence. The full-field propagation is shown in Supplementary Fig. 2a, while internal field distributions within the multilayer Bragg microcavity are detailed in logarithmic (Supplementary Fig. 2b) and linear (Supplementary Fig. 2c) scales. Crucially, these results demonstrate that optical energy does not undergo geometric 'inward' bending or negative refraction within the ultrathin stack. Instead, the Fabry–Pérot resonance mediates a nonlocal response where source wavefronts interfere constructively at the orthogonal projection of the source on the output plane. Within the lateral spreading length W centered at this projection, the incident wavefront maintains a constructive phase relationship due to its slow spatial phase variation. As the lateral position deviates from the center, however, rapid phase mismatch disrupts the interference condition, suppressing transmission and reflecting the off-resonance energy.

It should be noted that the constructive interference mentioned primarily refers to the propagation from the film's input aperture to the image plane. For an object at a finite distance, there is an additional phase difference from the object to different points on the input aperture. This object-side phase difference can be considered an aberration, as a planar film is effectively focused at infinity. However, the system's inherently extended depth of field mitigates the impact of this aberration, allowing it to form a clear image.

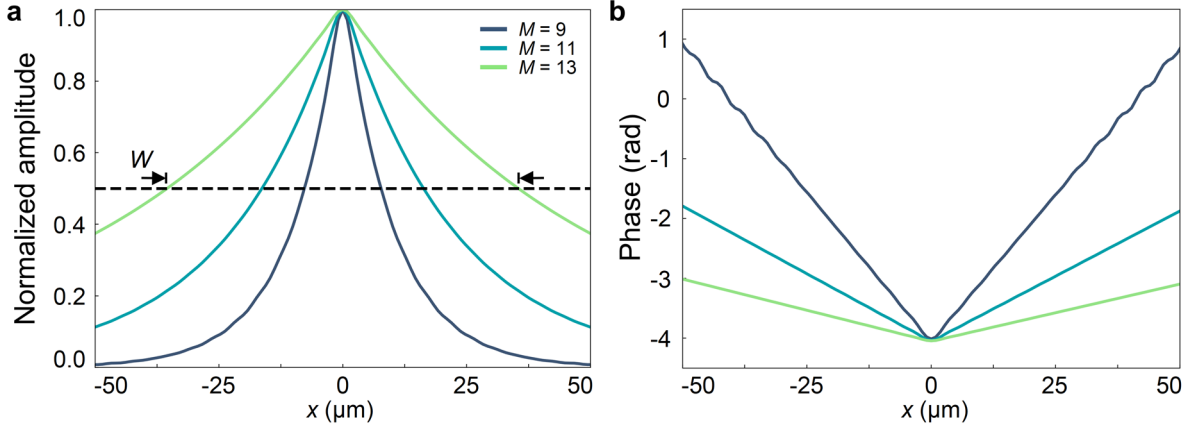
Extending the above one-dimensional analysis to two-dimensional scenarios, the normalized Fabry–Pérot multi-beam interference kernel $h(r)$ is given by,

$$h(r) = 2\pi f_0^2 K_0(2\pi \tilde{f}_0 r), \quad (\text{S4})$$

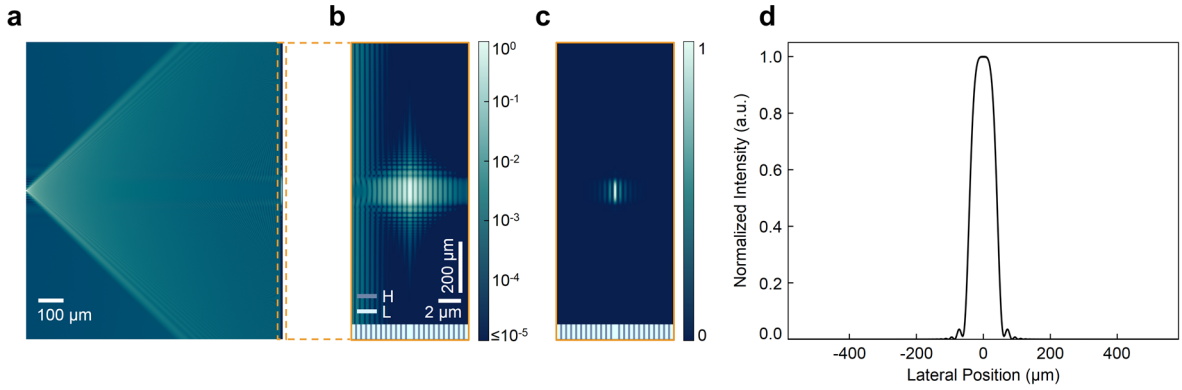
where K_0 denotes the zeroth-order modified Bessel function of the second kind, and $r = \sqrt{x^2 + y^2}$ represents the radial distance in the transverse plane. Integrating the kernel $h(r)$ along either the x - or y -axis reduces the two-dimensional form to the one-dimensional case described previously. This relationship can be explicitly expressed as,

$$\int_{-\infty}^{+\infty} K_0(2\pi\tilde{f}_0\sqrt{x^2+y^2})dy = \frac{1}{2\tilde{f}_0} e^{-2\pi\tilde{f}_0|x|}. \quad (\text{S5})$$

This integral demonstrates that the two-dimensional kernel degenerates to the same form as the one-dimensional case when projected along either lateral direction. Supplementary Figure 1 illustrates the evolution of the amplitude and phase of the integrated two-dimensional Fabry–Pérot multi-beam interference kernel along the y -direction, as the DBR pair number M increases.



Supplementary Figure 1 | Evolution of integrated Fabry–Pérot interference kernel along the y -direction with DBR pairs. **a**, The amplitude of the integrated kernel along the y -direction, where the lateral spreading length—defined as the full width at half maximum—increases from $14.7\ \mu\text{m}$ ($M=9$) to $31.1\ \mu\text{m}$ ($M=11$) and further to $69.6\ \mu\text{m}$ ($M=13$). **b**, The phase of integrated Fabry–Pérot interference kernel along the y -direction.



Supplementary Figure 2 | Electric field intensity distribution. **a**, Normalized electric field intensity distribution in logarithmic scale showing the propagation from a line source incident onto the nonlocal thin film. **b**, **c**, Normalized electric field intensity distribution inside the multilayer Bragg microcavity structure displayed in logarithmic (**b**) and linear (**c**) scales. The alternating bars at the bottom denote the layer structure, with H and L corresponding to high- and low-refractive-index materials, respectively. **d**, The corresponding normalized point spread function (PSF).

Supplementary Note 2. Lateral spreading length W and its scaling with the cavity Q-factor

The lateral spreading length W , defined as the full width at which the response amplitude falls to half of its maximum value, can be obtained directly from Eq. (S3) as,

$$W = \frac{\lambda \ln 2 n_H^M}{\sqrt{2\pi} n_L^{M+1.5}}. \quad (\text{S6})$$

W quantifies the spatial extent of the nonlocal coupling of object spatial frequencies enabled by the thin film, determining the achievable resolution. Meanwhile, the cavity's quality factor (Q-factor)—quantifying the resonance sharpness in wavelength—is determined from Eq. (S1) as,

$$Q = \frac{\pi \sqrt{\rho}}{1 - \rho}. \quad (\text{S7})$$

From Eqs. (S6) and (S7), the relationship between the cavity Q-factor and the lateral spreading length W is given by,

$$W = \frac{\sqrt{2} \ln 2 \lambda}{\pi n_L} \sqrt{Q}. \quad (\text{S8})$$

This result highlights that W increases with the cavity Q-factor. This relation underpins the resolution enhancement enabled by stronger nonlocal wavefront coupling through the Bragg microcavity.

Supplementary Note 3. Optical transfer function broadening through angular spectrum confinement with the nonlocal thin film

The angular spectrum corresponding to a point object propagating through free space can be represented as,

$$A(k_x, k_y) = \exp \left[i \frac{2\pi}{\lambda} \rho \sqrt{1 - \left(\frac{\lambda k_x}{2\pi} \right)^2 - \left(\frac{\lambda k_y}{2\pi} \right)^2} \right], \quad (\text{S9})$$

where ρ denotes the object distance from the imaging sensor, λ is the wavelength of illumination, and k_x, k_y are the components of the transverse wavevector $\mathbf{k}$, respectively. The optical transfer function (OTF) for free-space propagation is given by the autocorrelation of $A(\mathbf{k})$ shown in Supplementary Fig. 3a as,

$$\text{OTF}(\mathbf{k}) = A(\mathbf{k}) \otimes A^*(-\mathbf{k}), \quad (\text{S10})$$

where $\otimes$ denotes the convolution operation. Because $A(\mathbf{k})$ exhibits a chirped nature in free space, its autocorrelation is sharply peaked and closely approximates a delta function, indicating minimal overlap between spatial frequencies after propagation. This implies that free-space propagation alone does not allow interference among different spatial frequency components, thus inherently lacking imaging capability under incoherent illumination.

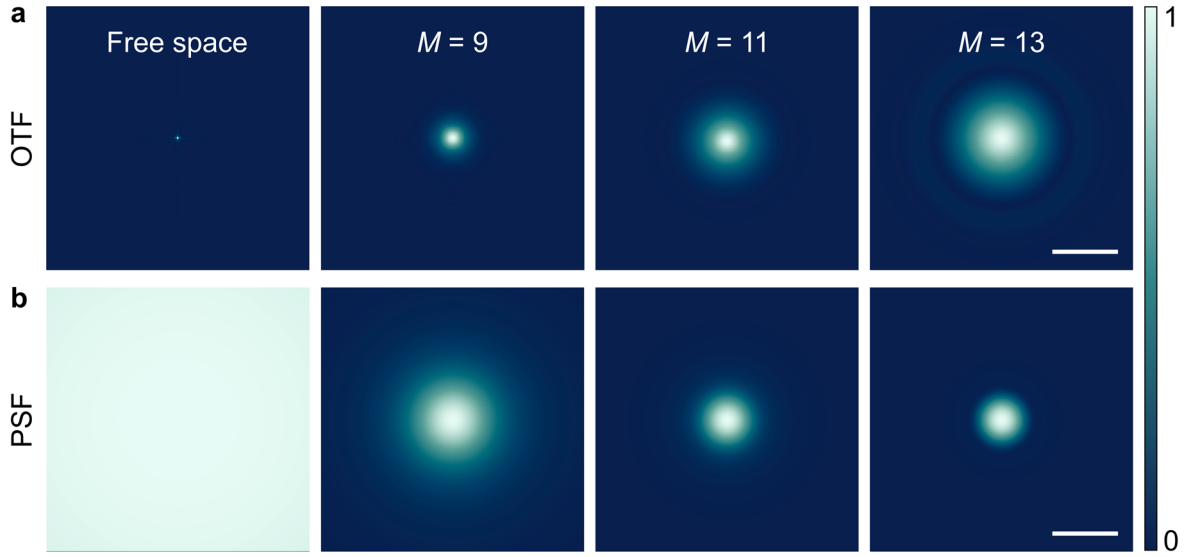
When the nonlocal thin film with transmission function $t(\mathbf{k})$ is introduced, as shown in Supplementary Fig. 3, the OTF is modified as,

$$\text{OTF}(\mathbf{k}) = [A(\mathbf{k})t(\mathbf{k})] \otimes [A^*(-\mathbf{k})t^*(-\mathbf{k})]. \quad (\text{S11})$$

This selective filtering confines the otherwise broad angular spectrum $A(\mathbf{k})$. The imposed confinement prevents destructive interference from widely diverging chirped wavefront components, thereby significantly enhancing the effective imaging aperture and broadening the OTF. Consequently, as demonstrated in Supplementary Fig. 3b, this broadened OTF leads directly to a narrower point spread function (PSF) through the inverse Fourier transform as,

$$\text{PSF} = \mathcal{F}^{-1}[\text{OTF}(\mathbf{k})] = |\mathcal{F}^{-1}[A(\mathbf{k})t(\mathbf{k})]|^2. \quad (\text{S12})$$

As the number of DBR pairs M increases, the angular selectivity of the nonlocal thin film is further enhanced, thereby broadening the OTF and correspondingly narrowing the PSF, leading to improved imaging resolution.



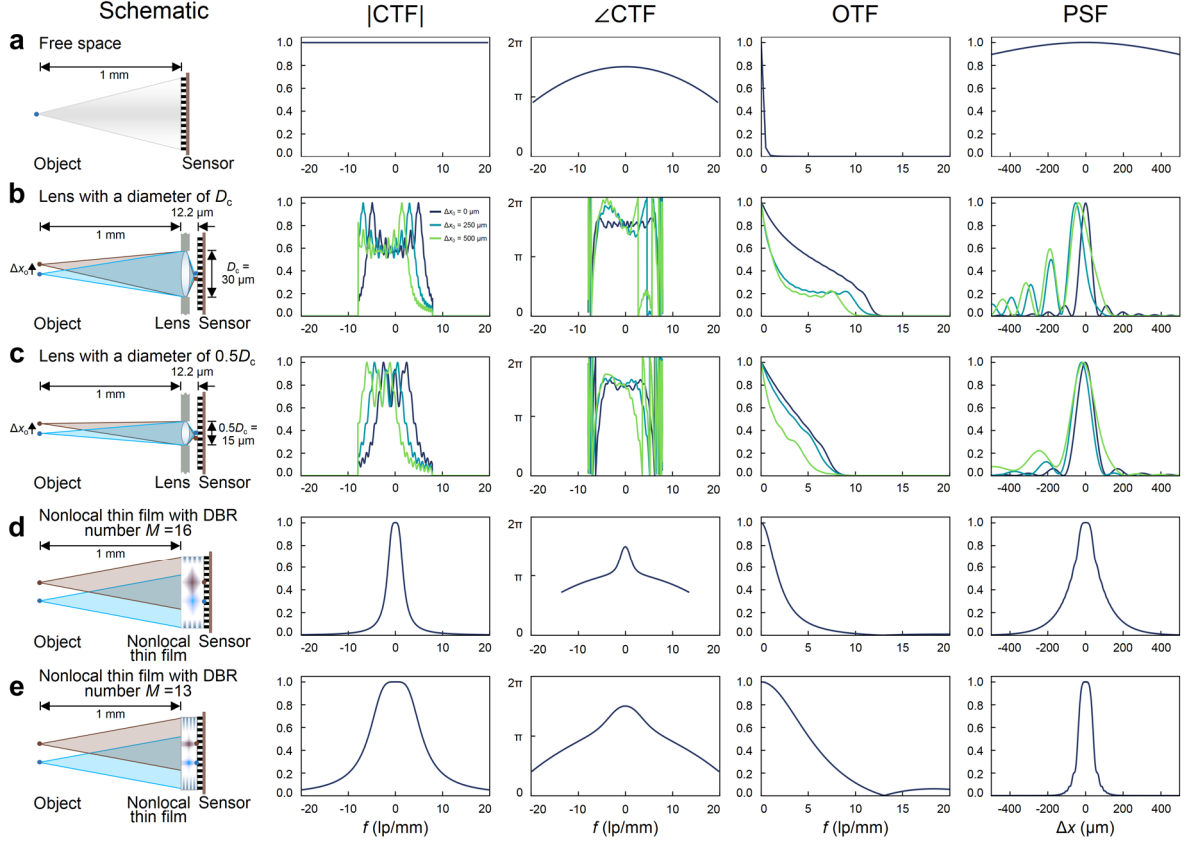
Supplementary Figure 3 | Optical transfer function (OTF) broadening and point spread function (PSF) narrowing through angular spectrum filtering by the nonlocal thin film. **a**, OTFs and **b**, corresponding PSFs for free-space propagation compared to nonlocal thin films with varying numbers of distributed Bragg reflector (DBR) pairs M , illustrating the broadening of the OTF and resultant narrowing of the PSF through selective filtering and coupling of spatial frequencies. Scale bars represent 12.5 lp/mm for the OTF and 150 μm for the PSF.

We clarify that while the topological difference allows us to bypass the thickness bound, the physical principle of imaging remains identical in both conventional translationally-variant imaging and our translationally-invariant nonlocal thin film imaging. To illustrate this, we analyze the object-space coherent transfer function (CTF) and its phase evolution, as illustrated in Supplementary Fig. 4. As quantified in Supplementary Fig. 4a, image formation is governed by the usable bandwidth of the object-space OTF, which is the autocorrelation of the CTF. While free-space propagation theoretically transmits a wide coherent band with nearly unit magnitude (Supplementary Fig. 4a, CTF amplitude panel), it inherently introduces a rapidly oscillating chirp phase (Supplementary Fig. 4a, CTF phase panel). This oscillation leads to destructive interference during the autocorrelation process, causing the OTF to collapse to a near-delta-like response with negligible bandwidth (Supplementary Fig. 4a, OTF amplitude panel). Consequently, some form of angular spectrum filtering (low-pass filtering) is generally

needed for image formation, to prevent the CTF phase from undergoing large variations, thereby broadening the effective OTF.

Conventional translationally-variant optics faces a field-dependent constraint in achieving this balance because all object points share a single pupil. For a conventional lens with a single shared aperture of diameter D_c (Supplementary Fig. 4b), although the coherent passband is wide (Supplementary Fig. 4b, CTF amplitude panel), where the CTF calculation accounts for the geometric magnification, different field points correspond to different chief-ray directions. Consequently, the CTF phase cannot maintain a small variation across the shared pupil for all object positions simultaneously, leading to strong phase oscillations and a rapid decay of OTF content for off-axis points (Supplementary Fig. 4b, CTF phase and OTF amplitude panels). To stabilize the OTF across the field, the aperture must be stopped down to $0.5D_c$ (Supplementary Fig. 4c). While this reduces phase variation, it imposes a global restriction on the spatial frequency support, resulting in reduced bandwidth and resolution for all points as the field of view (FOV) increases. This trade-off is manifested in the optimum aperture of conventional imaging systems, where the human eye attains its best optical quality at an intermediate pupil diameter² and a camera lens is sharpest at a moderate f -number³⁻⁵.

In contrast, our translationally-invariant thin-film architecture effectively assigns an individual aperture to each object point. This bypasses the single-aperture assumption, allowing the system to maintain a small CTF phase variation for all object positions without the constraints of a global pupil geometry. However, we note that the filtering strength, which determines the angular bandwidth of the transmitted spectrum, must be carefully optimized. If the angular passband is narrowed too aggressively, as shown in Supplementary Fig. 4d, the transmitted spatial-frequency support becomes overly restricted. While the phase variation is small across the passband (Supplementary Fig. 4d, CTF phase panel), the OTF bandwidth shrinks (Supplementary Fig. 4d, OTF panel) and the PSF broadens (Supplementary Fig. 4d, PSF panel). Thus, our design achieves the optimal balance shown in Supplementary Fig. 4e. With the nonlocal Bragg microcavity film operating with an appropriate angle-dependent transmission, the angular spectrum is confined such that the CTF phase variation remains small across the passband (Supplementary Fig. 4e, CTF phase panel), yielding a substantial OTF bandwidth (Supplementary Fig. 4e, OTF panel) enabling high-resolution imaging in an ultrathin form factor.



Supplementary Figure 4 | Comparison of coherent and incoherent transfer metrics. The results are obtained from 1D simulations. **a**, Free-space propagation. **b**, **c**, A conventional lens system with a single shared aperture of diameter D_c (**b**) and $0.5D_c$ (**c**), where the image distance is set to $12.2 \mu\text{m}$. **d**, **e**, The proposed nonlocal thin film with strong angular selectivity (DBR pair number $M = 16$) (**d**) and moderate angular selectivity ($M = 13$) (**e**). From left to right, columns display the system schematic, normalized CTF amplitude, CTF phase, OTF amplitude, and object-space PSF (mapped from the image plane via magnification). For **b**, **c**, results are plotted for three representative object lateral positions Δx_0 (defined as the lateral displacement from the optical axis) of $0, 250$, and $500 \mu\text{m}$.

Supplementary Note 4. Thickness bound of translationally-invariant imaging systems

As is pointed out by D. A. B. Miller⁶, for a one-dimensional translationally-variant imaging system, whose input is a lens surface and output is an array of N pixels, the wavefront produced by each object point must be distinguished across a common, finite aperture. The sampling theorem equates the N resolvable input sample points with N orthogonal input fields on that aperture. An intermediate surface introduced mathematically to divide both input and image planes into two halves gives an overlapping nonlocality of $N/2$. This leads directly to a minimum total track length of $N\lambda/4n_{\text{max}}$, where n_{max} is the maximum refractive index between the input plane and the image plane.

By contrast, the imaging system presented in this study is translationally-invariant. In a translationally-invariant imaging system, each object position corresponds to an effective aperture that varies with the position. Consequently, no single physical aperture is filled by wavefronts emitted by all of the object points simultaneously. Because the prerequisite

common aperture is absent, the overlapping nonlocality is no longer constrained to $N/2$ and the thickness bound for conventional translationally-variant imaging systems proposed in Ref. [10] of the main text no longer applies.

Besides, for translationally-invariant imaging systems, the response received by each pixel is formed by the integral of the product of the wavefront emitted by an object point and the nonlocal response kernel. Consequently, the imaging cutoff frequency is governed by the lateral spreading of the nonlocal response kernel. For the nonlocal response kernel, scaling its size alters the lateral spreading and thus the imaging cutoff frequency. However, this scaling does not change the number of significant singular values⁶, and therefore the thickness bound remains unaffected. Moreover, extending the FOV with an increasing number of pixels does not require altering the imaging kernel, meaning the system's thickness remains independent of the pixel count. Therefore, from the non-discrete perspective, spatially translationally-invariant imaging systems are not constrained by thickness bound imposed by the number of nonlocal response's significant singular values.

This result is consistent with the theoretical analysis based on singular value decomposition (SVD). For a 1D nonlocal thin film, the nonlocal response kernel takes the form of a complex exponential function, $h(x) \propto e^{-a|x|}$. Consequently, the corresponding submatrix can be expressed as,

$$D_{\text{RL}} = \begin{bmatrix} e^{-3a\Delta} & e^{-4a\Delta} & e^{-5a\Delta} \\ e^{-2a\Delta} & e^{-3a\Delta} & e^{-4a\Delta} \\ e^{-a\Delta} & e^{-2a\Delta} & e^{-3a\Delta} \end{bmatrix}. \quad (\text{S13})$$

Mathematically, this is a Toeplitz matrix with a rank of exactly 1. This implies that the number of non-zero singular values is also 1, regardless of the number of terms included or the sampling interval Δ . This yields the overlapping nonlocality of $C = 2$. This result indicates that the overlapping nonlocality is indeed minimal, and the fundamental thickness required for this specific imaging modality can approach the half-wavelength scale. This confirms that the imaging bound for our specific transfer function is indeed minimal and fundamentally decoupled from the pixel count, substantiating our claim of bypassing the conventional thickness bound.

To further illuminate the nature of optical interactions in our system, we clarify that by “interactions,” we refer to the linear coupling coefficients mapping the optical field from spatial sampling points on the input aperture to sampling points on the image plane, corresponding to the values of the nonlocal response kernel. In the matrix formalism, these interactions constitute the matrix elements of the transmission matrix connecting the input and output degrees of freedom. Our statement regarding orthogonality addresses the difference in the number of significant singular values of this transmission matrix between conventional imaging systems and our translationally-invariant nonlocal film.

We first consider the “interactions” within a conventional imaging system consisting of a thin lens of focal length f followed by free-space propagation over a distance $z = f$. The relationship between the input field $E_{\text{in}}(x)$ and the output field $E_{\text{out}}(u)$ is governed by the Fresnel diffraction integral multiplied by the lens's phase transmission function. Specifically, the lens introduces a quadratic phase factor $\exp[-ikx^2/(2f)]$, which cancels the quadratic phase term inherent to the Fresnel propagation kernel. Consequently, the output field relates to the input field via a Fourier transform integral $E_{\text{out}}(u) \propto \int E_{\text{in}}(x) \exp\left[-i\frac{k}{f}ux\right] dx$. To analyze the system's

capacity to resolve N distinct pixels, we discretize this continuous relationship. Ideally, the system is sampled at intervals Δx and Δu satisfying the sampling theorem, such that the phase factor $\frac{k}{f}\Delta u\Delta x = \frac{2\pi}{N}$. For a simplified 4-point model ($N=4$) with indices $m, n \in \{0, 1, 2, 3\}$, this discretization yields a transmission matrix T_c with elements given by $T_{cmn} = \exp\left[-i\frac{k}{f}(m\Delta u)(n\Delta x)\right]$. Substituting the condition, the matrix elements become $\exp\left[-i\frac{\pi}{2}mn\right]$,

$$T_c = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -i & -1 & i \\ 1 & -1 & 1 & -1 \\ 1 & i & -1 & -i \end{pmatrix}. \quad (\text{S14})$$

Following the block-matrix analysis, we partition the system into left (indices 0, 1) and right (indices 2, 3) halves. The submatrix T_{RL} , representing the “interactions” crossing the dividing surface from the right input to the left output, is,

$$T_{RL} = \begin{pmatrix} 1 & 1 \\ -1 & i \end{pmatrix}. \quad (\text{S15})$$

The determinant of this block is non-zero ($\det(T_{RL}) = 1 + i$), indicating that it is full rank. This mathematical result confirms that in conventional imaging, the optical channels crossing the dividing surface are linearly independent. It is this requirement for maintaining full rank across the aperture that necessitates a system thickness scaling linearly with the number of pixels.

In stark contrast, our translationally-invariant system is governed by a nonlocal exponential decay kernel $h(x) \propto e^{-\alpha|x|}$. When discretized for the same 4-point configuration with spacing Δ , the transmission matrix T_{film} becomes a Toeplitz matrix where elements depend only on the distance $|m - n|$,

$$T_{\text{film}} = \begin{pmatrix} 1 & e^{-a\Delta} & e^{-2a\Delta} & e^{-3a\Delta} \\ e^{-a\Delta} & 1 & e^{-a\Delta} & e^{-2a\Delta} \\ e^{-2a\Delta} & e^{-a\Delta} & 1 & e^{-a\Delta} \\ e^{-3a\Delta} & e^{-2a\Delta} & e^{-a\Delta} & 1 \end{pmatrix}. \quad (\text{S16})$$

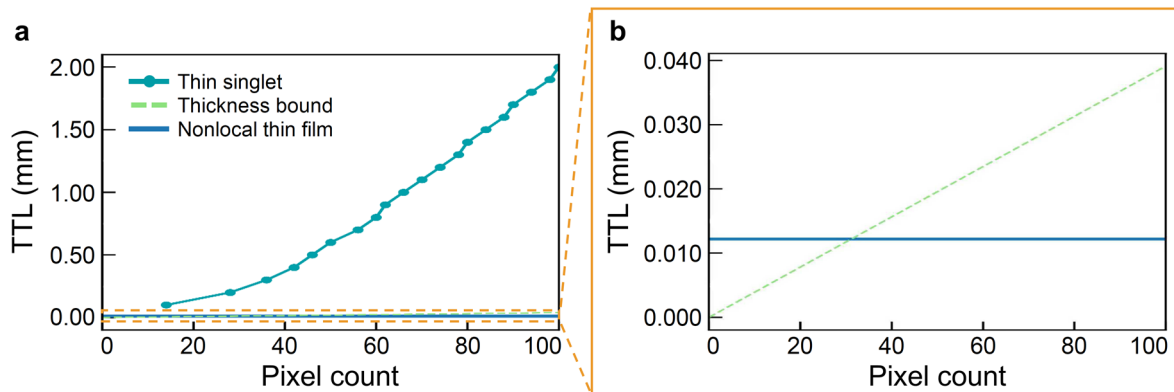
Examining the same cross-section submatrix T_{RL} (Right input to Left output), we obtain,

$$T_{RL} = \begin{pmatrix} e^{-2a\Delta} & e^{-3a\Delta} \\ e^{-a\Delta} & e^{-2a\Delta} \end{pmatrix}. \quad (\text{S17})$$

Here, the first row is simply the second row multiplied by a scalar $e^{-a\Delta}$, resulting in a vanishing determinant and a rank of 1. This demonstrates that the “interactions” crossing the interface in our system are not mutually orthogonal. Rather, they are highly correlated and linearly dependent. Because the rank of the cross-section matrix is determined by the kernel’s decay profile rather than the number of pixels, the overlapping nonlocality does not scale with the pixel count. This lack of orthogonality allows our device to operate well below the thickness bound derived for conventional systems.

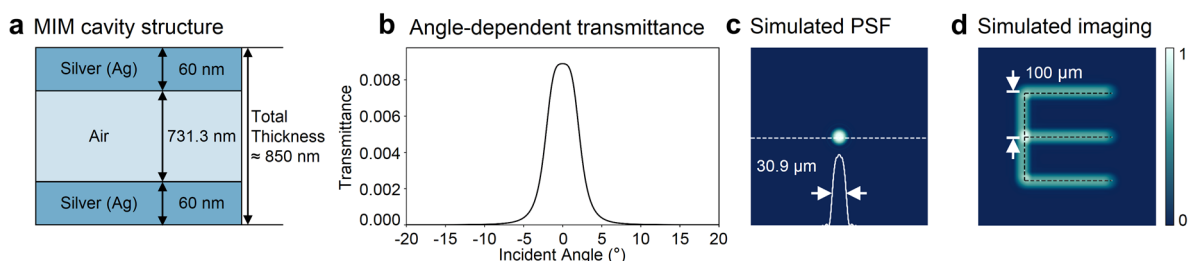
To quantitatively validate this difference in scaling laws, we simulated the relationship between the total track length (TTL) and the pixel count for a thin singlet, as illustrated in Supplementary Fig. 5. Here, the pixel count is determined by the field of view within which

the PSF maintains a single peak. The simulation results show that the thin singlet exhibits a monotonic scaling relationship between TTL and pixel count, confirming that translationally-variant systems must proportionally increase thickness to accommodate the overlapping nonlocality. In stark contrast, the nonlocal thin film maintains a constant TTL of $12.2\ \mu\text{m}$ independent of the pixel count, demonstrating that our translationally-invariant architecture decouples system thickness from the pixel count, thereby bypassing the thickness constraints inherent to conventional single-aperture topologies.



Supplementary Figure 5 | Scaling behavior of TTL versus pixel count. **a**, Lines represent the relationship between TTL and pixel count for the thin singlet, thickness bound, and the nonlocal thin film, respectively. **b**, Zoomed-in view of a partial region of (a).

To probe the ultimate physical limit of this architectural paradigm, we simulated a metal-insulator-metal (MIM) Fabry-Pérot cavity (Supplementary Fig. 6a). By replacing the distributed Bragg reflectors (DBRs) with silver layers, we dramatically reduced the total device thickness. The simulation demonstrates that even at a thickness of approximately 850 nm (approaching the half-wavelength scale at 1550 nm), the system maintains the necessary angularly selective transmittance (Supplementary Fig. 6b). It yields a PSF with a full-width-at-half-maximum (FWHM) of $30.9\ \mu\text{m}$ at an object distance of $150\ \mu\text{m}$ (Supplementary Fig. 6c), sufficient to clearly resolve a letter “E” target with $100\ \mu\text{m}$ line spacing (Supplementary Fig. 6d). This result confirms that our architecture enables imaging at wavelength-scale thicknesses, a regime fundamentally unattainable for conventional systems restricted by the thickness bound.



Supplementary Figure 6 | Validation of imaging with a half-wavelength-thick metal-insulator-metal (MIM) film. **a**, Schematic of the MIM Fabry-Pérot cavity, with a total thickness of around 850 nm. **b**, The corresponding angle-dependent transmittance. **c**, Simulated PSF at an object distance of $\rho = 150\ \mu\text{m}$ under $\lambda = 1550\ \text{nm}$ illumination, showing a FWHM of $30.9\ \mu\text{m}$. **d**, Simulated imaging of a letter “E” with a line spacing of $100\ \mu\text{m}$.

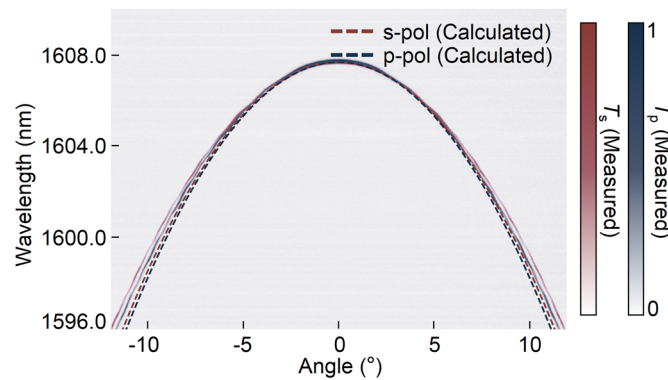
We further evaluated alternative nonconventional architectures with low overlapping nonlocality, specifically microlens array⁷ and contact imaging based on gradient-index (GRIN) lens array⁸. While these approaches reduce overlapping nonlocality by distributing apertures, they face significant practical limitations for deterministic continuous imaging at a finite object standoff.

In a microlens array, apertures cannot overlap continuously, which imposes a geometric tradeoff between lenslet pitch and aperture size and can lead to nonuniform sampling. Although one might intend to utilize only the central pixel behind each microlens, standard microlenses inherently possess a certain FOV even if it remains unused. Conversely, designing a lens element specifically to resolve only one or two pixels, which would be necessary to minimize thickness to the extreme, is nontrivial. Since individual GRIN lenses cannot spatially overlap, they must possess a certain FOV to avoid dead zones. Consequently, the thickness is inevitably dictated by the required imaging FOV for each lens, and such systems also suffer from crosstalk between adjacent elements arising from large incident angles.

Therefore, in practice, to achieve deterministic continuous imaging, to the best of our knowledge, there is no alternative imaging approach meeting the requirements other than conventional single aperture imaging systems. In contrast, as a new nonconventional imaging approach, our thin film architecture allows the effective apertures corresponding to adjacent pixels to be continuously distributed. This enables deterministic imaging in an ultrathin form factor while bypassing the thickness bound inherent to conventional single aperture optics.

Supplementary Note 5. Polarization-insensitive response at near-normal incidence

Supplementary Figure 7 compares theoretically calculated and measured angle-resolved transmittance spectra of the fabricated nonlocal thin film, displaying results for both *s*- and *p*-polarizations simultaneously within the same figure. Experimental data, represented by red and blue colors corresponding to *s*- and *p*-polarizations respectively, were acquired by rotating the polarization axis of a polarizer during characterization. Dashed lines indicate the theoretical resonance angles as a function of wavelength, calculated using the transfer matrix method and shifted in central wavelength to align with measured resonances, exhibiting good agreement with experimental results. Supplementary Figure 7 demonstrates the negligible difference between *s*- and *p*-polarized transmittance at near-normal incidence. Such polarization-insensitive behavior originates from the rotationally-invariant structure of the thin film.



Supplementary Figure 7 | Comparison between theoretically calculated and measured angle-resolved transmittance spectra of the fabricated nonlocal thin film. Experimental and theoretical angle-resolved transmittance spectra of the nonlocal thin film under *s*- (red) and *p*- (blue) polarization. Dashed lines indicate the theoretical resonance angles as a function of wavelength.

Supplementary Note 6. Axial placement flexibility of nonlocal thin film through transfer function commutativity

As shown in Eq. (S9), since $A(\mathbf{k})$ includes the phase factor $\exp[ik_z\rho]$ with $k_z = 2\pi/\lambda \sqrt{1 - (\lambda k_x/2\pi)^2 - (\lambda k_y/2\pi)^2}$, one may regard ρ as the sum of ρ_1 (from the object to the thin film) and ρ_2 (from the film to the detector, including protective spacer). In the k -space, the wavefront first acquires phase factor $\exp[ik_z\rho_1]$, then is filtered by $t(\mathbf{k})$, and finally picks up another phase factor $\exp[ik_z\rho_2]$. Since $\exp[ik_z\rho_1]$, $t(\mathbf{k})$ and $\exp[ik_z\rho_2]$ are all multiplicative (diagonal) operators in k -space, they can commute, so one can combine them as,

$$t(\mathbf{k}) \exp[ik_z(\rho_1 + \rho_2)] = A(\mathbf{k}) t(\mathbf{k}), \quad (\text{S18})$$

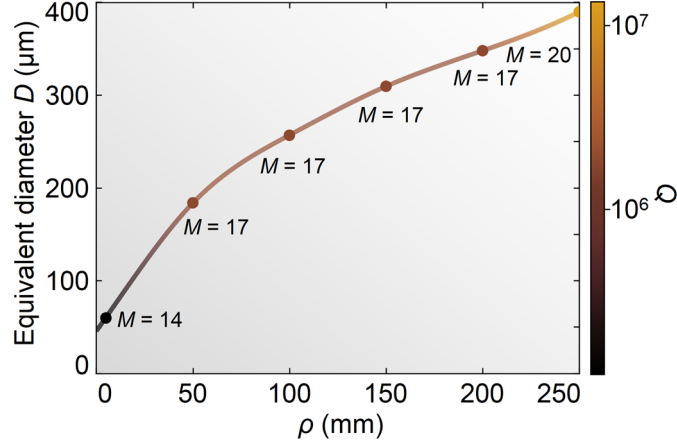
without changing the PSF. Hence the PSF depends only on the total propagation distance ρ , making the imaging system insensitive to the film's axial placement.

Supplementary Note 7. 2D imaging resolution at different object distances

Because imaging using the nonlocal thin film is inherently telecentric, the system's magnification remains fixed at unity, and spatial resolution is thus determined solely by the PSF's FWHM. For finite object distances, the PSF is obtained by calculating the intensity distribution on the image plane of a spherical wave emitted by an object point after filtering by the thin film, as detailed in Supplementary Note 3. In the limit of infinite object distance, the PSF in the angular domain follows directly from the thin film's transmission function.

As shown in Supplementary Fig. 3b and Fig. 2d, the FWHM of the PSF at finite object distances generally decreases with an increasing number of DBR pairs M and Q-factor, indicating enhanced resolution. This improvement in resolution arises because the multilayer Fabry–Pérot cavity structure enhances angular selectivity, allowing a broader lateral extent of optical fields to interfere constructively. Consequently, a wider range of spatial frequencies contributes effectively to image formation, thereby improving the resolution. When the Q-factor is very high, the FWHM slightly increases with the increasing Q-factor due to the residual quadratic phase carried by the spherical wavefront within the aperture. This effect disappears as the object distance increases.

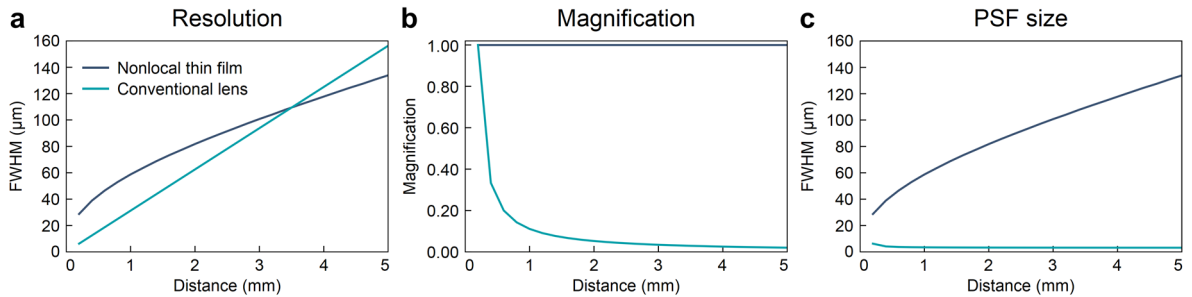
Supplementary Figure 8 shows the equivalent diameter D —defined as the diameter of an ideal on-axis lens producing the same resolution—as a function of object distance, where each point represents the highest resolution achieved by optimizing the DBR pair number M for that specific distance. As object distance increases, the equivalent diameter D correspondingly rises, and in the limit of infinite object distance, the incoming wavefront's quadratic phase vanishes so that D becomes a monotonically increasing function of the Q-factor, consistent with the trend in Fig. 2c of the main text.



Supplementary Figure 8 | Optimized equivalent diameter dependence on the thin film's Q-factor and object distance ρ .

To further clarify the trade-off between object distance and imaging performance, we emphasize that the degradation of resolution with distance is a fundamental physical phenomenon inherent to both conventional lens-based systems and our nonlocal thin film, although arising from different mechanisms. As illustrated in Supplementary Fig. 9a, the resolution of our nonlocal thin film (dark blue line) degrades with increasing distance, a trend that is similar to the fitted resolution scaling of a conventional lens with a diameter of 50 μm (teal line).

In conventional lens systems, resolution degrades primarily due to the reduction in magnification at longer distances, as the object-side resolution is derived by dividing the image-side PSF FWHM by the magnification. Supplementary Fig. 9b shows that a conventional lens suffers from a drop in magnification as the object moves away, whereas our system maintains constant unity magnification owing to its telecentric nature. As shown in Supplementary Fig. 9c, the spatial extent of the PSF in our system scales with the propagation distance, leading to a similar trend of reduced resolution for distant objects. Nevertheless, the current implementation is strategically optimized for close-range applications where physical space is critically constrained. In domains such as medical endoscopy⁹, wearable sensors¹⁰, and robotic tactile sensing^{11, 12}, the strict spatial limitations often dictate that form factor takes precedence over ultimate imaging resolution, necessitating the extreme compactness that our system provides. In these contexts, the ultracompact form factor enabled by our architecture provides a capability that bulky conventional optics cannot achieve.



Supplementary Figure 9 | Comparison of imaging characteristics over distance. a,

Resolution defined by FWHM (calculated as the image-side PSF FWHM divided by magnification) as a function of object distance. The dark blue line represents the nonlocal thin film, while the teal line indicates the fitted trend for a conventional lens with a diameter of 50 μm . **b**, Magnification as a function of object distance. **c**, PSF FWHM on the image plane as a function of object distance.

Supplementary Note 8. Comprehensive performance comparison across imaging modalities

To evaluate the imaging performance of the proposed nonlocal thin film-based system, we quantitatively benchmarked it against conventional singlet refractive lens-based imaging systems as well as metalens-assisted, space plate-assisted, and computational lensless imaging systems. For proper benchmarking across all configurations, we assumed an identical sensor diameter of 4.18 mm (same as the sensor diameter used in the main text) and an identical object-to-sensor distance of 2.25 mm, which is the minimum achievable distance limited by the protective cover glass of the CMOS sensor and corresponds to the configuration used in Figs. 1 and 2 of the main text. The target absolute magnification was set to 1:1. For conventional spatially-variant imaging systems, the resolution varies across the FOV.

For the conventional lens, we optimized a double-spherical singlet in Zemax OpticStudio with the entrance pupil placed at the first surface and a fixed TTL of 1.125 mm. In the optimization, the radii of curvature of both surfaces and the entrance-pupil diameter were treated as free variables. The merit function was constructed to minimize aberrations across the full FOV, suppress distortion, and promote uniform image quality across field positions. The optimizer thus balanced off-axis aberrations against numerical aperture, and the final design was chosen to minimize the FWHM of the PSFs. The final lens design has an entrance-pupil diameter of 0.15 mm, convex first and second surfaces with radii of curvature of 3.03 mm and 0.39 mm, respectively, and a center thickness of 0.10 mm.

In the case of the metalens-assisted imaging systems, the planar metalens features quadratic phase modulation. Because a single planar element inherently lacks sufficient degrees of freedom to suppress off-axis aberrations, especially field curvature and distortion, the PSFs exhibit larger FWHM than those of the optimized conventional refractive singlet. We optimized the aperture diameter to 60 μm to minimize the PSF FWHM across the FOV, balancing on-axis resolution against off-axis aberrations.

For space plate-assisted imaging systems, the TTL was set to 0.366 mm to maintain consistent object-to-image plane distances across configurations. Similar to the metalens-assisted imaging system, the aperture diameter was optimized to 92 μm , balancing on-axis resolution and off-axis aberrations to minimize the PSF FWHM. The space-plate compression ratio is set to 5.15 according to the typical values reported¹³, and the maximum-compression configuration is adopted. Notably, bandwidth constraints yield an actual FOV smaller than the designed value, while non-ideal transfer functions of the multilayer film further degrade resolution.

For computational lensless imaging at small TTL, amplitude mask modulation is mainly modeled as shadowing effects¹⁴. A pinhole is the simplest amplitude mask. Shadow-based coded aperture systems using binary masks can be regarded as combinations of pinholes¹⁴⁻¹⁶. The image formed on the sensor can be viewed as a superposition of many pinhole images. The

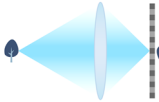
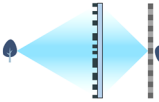
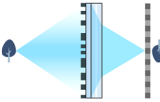
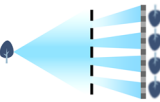
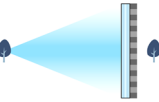
relationships among linewidth, resolution, and off-axis aberration are therefore similar to those of pinhole imaging. We therefore analyze a single pinhole for comparison. Considering the effects of pinhole size on both on-axis and off-axis PSFs, we optimized the pinhole diameter to $108\text{ }\mu\text{m}$ to make the PSF FWHM at the field center and at the field edge comparable, at which point the PSF FWHM within the FOV attains its minimum. Moreover, reducing TTL decreases magnification, which tightens the required sensor pixel size, and increases image-side aberration, which reduces reconstruction quality.

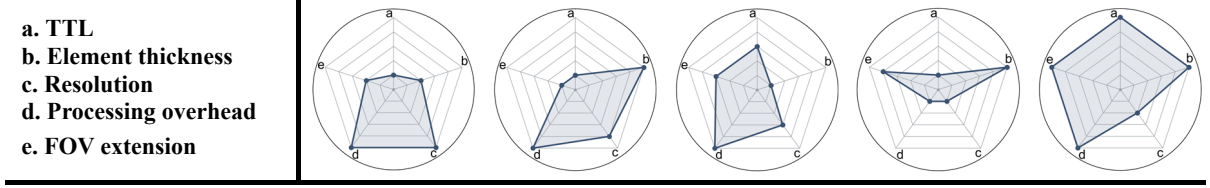
In this work, we use the same comparison configuration as in the main text with an object-to-sensor distance of 2.25 mm . For the DBR pair number $M = 13$, the corresponding simulation yields a PSF FWHM of $86.8\text{ }\mu\text{m}$ at this distance and an absolute magnification of 1. Besides, the FOV can be readily expanded by increasing the sensor area.

Supplementary Table 1 summarizes the TTL, resolution, and processing overhead as well as schematic illustrations for each modality, demonstrating the competitive performance of the nonlocal thin film-based imaging system. A comprehensive performance comparison is visualized using polar plots, where a larger radial value consistently denotes superior performance. The TTL mentioned here is determined for each system under the aforementioned magnification and object-to-sensor distance. The resolution refers to the on-axis resolution. For the spatially-variant imaging systems, the FOV extension is generally inversely correlated with the off-axis resolution.

Based on this performance profile, the current implementation is not intended to serve as a direct replacement for microscope objectives, but rather represents a strategic trade-off where extreme compactness is prioritized to enable image formation within a lensless architecture. In fact, not all imaging systems target half-wavelength resolution as the primary objective, and thus a high numerical aperture is not a universal requirement. For specific applications such as medical endoscopy⁹, wearable sensors¹⁰, and robotic tactile sensing^{11, 12}, the priority is often the extreme compactness required to integrate sensors into confined spaces rather than imaging resolution. In these specific practical scenarios, the thin film can be attached directly onto the image sensor and operated under active illumination, enabling lensless image formation with an ultracompact form factor unattainable with bulky conventional optics. Furthermore, the importance of bypassing the imaging thickness bound remains valid regardless of the numerical aperture.

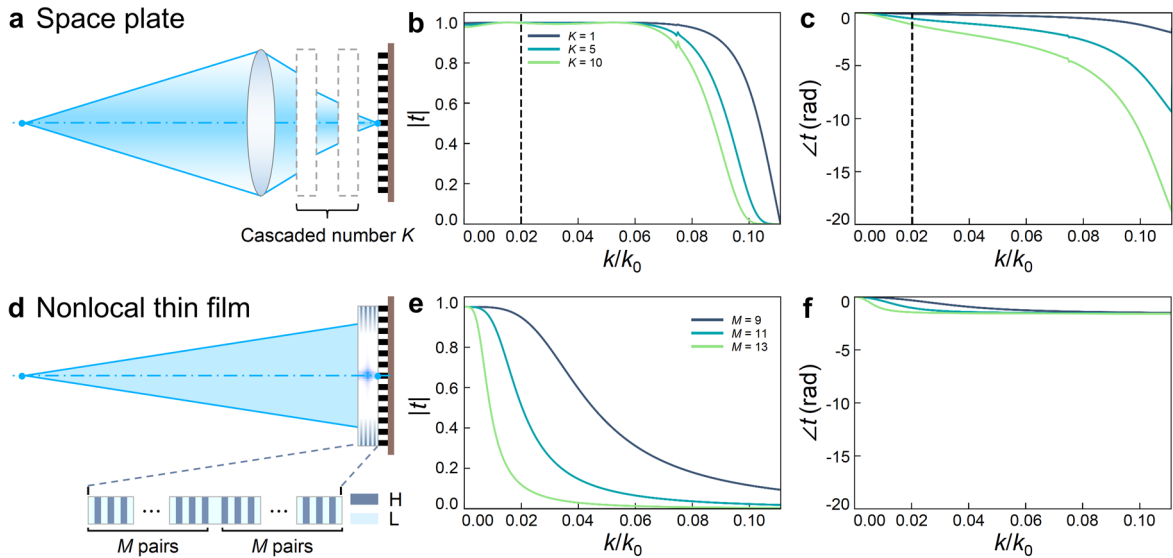
Supplementary Table 1. Quantitative performance comparison across imaging modalities.

		Singlet refractive lens	Metalens	Space plate	Computational lensless imaging	This work
Schematic illustration						
TTL (mm)		1.125	1.125	0.366	1.125	0.0122
Resolution (μm)	On axis	12.6	30.5	33.3	115.4	86.8
	Off axis	184.6	309.8	135.1	130.3	86.8
Processing overhead		No	No	No	Yes	No



To further elucidate the fundamental difference between our approach and the space-plate-based systems included in the comparison, we note that while both utilize resonant structures, the functionality of our proposed nonlocal thin film is fundamentally distinct from that of a space plate. The critical difference lies in the distinct roles the resonances play in the image formation process and the resulting system architectures.

Specifically, the space-compression functionality described in Ref. [11] of the main text relies on emulating the quadratic phase transfer function of free space over a specific angular range. As illustrated in Supplementary Fig. 10a, these space plates are typically integrated with a conventional lens to form an imaging system, serving to reduce the physical track length but retaining the translationally-variant nature of lens-based optics. Supplementary Figures 10b and 10c show the amplitude and phase of the angular transfer functions corresponding to the parameters in Ref. [11]. Increasing the number of cascaded layers K steepens the quadratic phase curvature to synthesize a large effective propagation length, while the effective working FOV remains limited by the effective working angular range (indicated by the dashed lines). Conversely, our approach does not rely on the strategy of compressing the optical path. Instead, as illustrated in Supplementary Fig. 10d, we utilize high-Q resonances to strictly confine the angular spectrum. Supplementary Figures 10e and 10f display the amplitude and phase of the angular transfer functions of the proposed nonlocal thin film. Increasing the number of DBR pairs M narrows the angular transmission bandwidth to directly shape the spatial impulse response for imaging. This angular filtering mechanism narrows the CTF and effectively broadens the OTF. Consequently, it produces an extended lateral spreading length W , which effectively forms a narrow point spread function to directly define the resolution of our translationally-invariant imaging system.



Supplementary Figure 10 | Comparison of the angular transfer functions between cascaded space plates and the proposed nonlocal thin film. a, Schematic of space plate-assisted lens imaging with varying numbers of cascaded layers K . b, c, Simulated amplitude

(b) and phase (c) responses of the space plates with $K = 1, 5, 10$, where the dashed lines indicate the effective working angular range. **d**, Schematic of the proposed nonlocal thin film imaging. **e, f**, Simulated amplitude (e) and phase (f) responses of the nonlocal thin film at the center wavelength with varying DBR pairs $M = 9, 11, 13$.

Furthermore, this distinction leads to fundamentally different system-level implications regarding the imaging thickness bound. As analyzed by Miller⁶, space plates cannot fundamentally break the imaging thickness limit. This is because a space-plate-based system typically relies on a shared physical aperture to resolve spatially distinct points; consequently, the system thickness remains fundamentally constrained by the overlapping nonlocality required to resolve a given number of pixels through that shared aperture. Conversely, our nonlocal thin film operates as a translationally-invariant system where an effective aperture is assigned individually to each object point via the lateral spreading mechanism. This architecture bypasses the single-aperture constraint that dictates the thickness limit in space-plate-based or conventional lens-based systems, enabling an ultrathin form factor that is unattainable by simply compressing free space.

Supplementary Note 9. Volumetric channel density analysis

To further evaluate the information capacity per unit volume, we adopt the volumetric channel density (VCD) metric¹⁷. We first compute the VCD for our demonstrated nonlocal thin film. For a consistent comparison, we model the image sensor as a circle with a diameter equal to the diagonal of the actual sensor used, resulting in an effective sensor area of 13.76 mm². Because the system is translationally-invariant, its imaging performance can be characterized by a single representative PSF. Using the theoretical PSF FWHM of 86.8 μm reported in the main text, we estimate the number of resolvable spots N_{spots} across the sensor area. To align with the scaling used in the reference literature, we approximate the spot area as πFWHM^2 , which yields approximately 581 resolvable spots on our 13.76 mm² sensor. With a film thickness of 12.2 μm , the total optical volume is approximately 0.168 mm³. To ensure scale independence, the volume is raised to the 2/3 power to match the dimensionality of the information capacity,

$$\text{VCD} = N_{\text{spots}}/V^{2/3}. \quad (\text{S19})$$

We find that our nonlocal thin film achieves a VCD of approximately 1910 spots/mm². To provide a benchmark, we next evaluate the VCD of a conventional imaging system using the idealized thin-lens reference model. To ensure a consistent comparison, we adapt the reference model to utilize the FWHM as the resolution metric. In this model, the number of resolvable spots is,

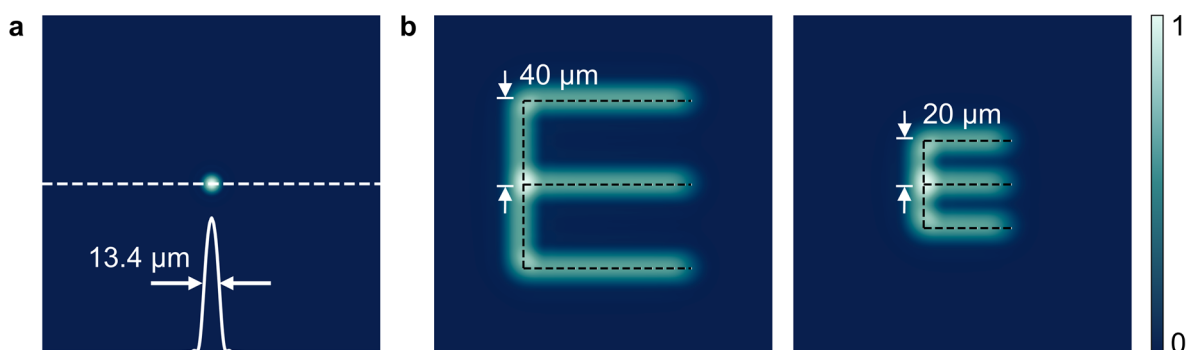
$$N_{\text{spots}} = \left(\frac{D}{\lambda}\right)^2 \frac{1 - \cos^4 \theta_{\text{max}}}{2}, \quad (\text{S20})$$

where θ_{max} is the maximum field half-angle, which is approximately 61.7° for the specified parameters. The optical volume is taken as that of a truncated cone between the lens and the image plane. We select an aperture diameter $D = 60 \mu\text{m}$ consistent with the thin singlet benchmark in Supplementary Note 8, where this value was optimized to minimize the PSF's FWHM across the FOV. We calculate a VCD of approximately 232 spots/mm² for the conventional thin-lens system.

Therefore, the VCD of our nonlocal thin film exceeds that of a conventional thin-lens system with comparable specifications by approximately a factor of 8.2. This comparison demonstrates that our nonlocal thin film not only bypasses the conventional imaging thickness bound but also achieves an advantage in volumetric channel density compared to traditional lens-based systems.

Supplementary Note 10. Strategies for resolution enhancement

As outlined in the main text, decreasing the object-to-sensor distance ρ or shifting the operational wavelength toward shorter values provides a simple yet effective route to improved resolution, achieved by proportionally scaling the layer thicknesses of the thin film. To illustrate this strategy, we simulated imaging performance at a wavelength of $\lambda = 532$ nm, with an object distance set to $\rho = 160$ μm . The number of DBR pairs is optimized to $M = 11$ for the best resolution at this distance. The resulting simulated PSF exhibits a FWHM of 13.4 μm as shown in Supplementary Fig. 11a. This configuration allows clear imaging of finer structures, exemplified by simulated imaging of letter “E” patterns with 40 μm and 20 μm line spacing as shown in Supplementary Fig. 11b. These results demonstrate that straightforward wavelength scaling and reduced object distance significantly improve resolution.



Supplementary Figure 11 | Resolution enhancement at shorter wavelength and reduced object distance. **a**, Simulated PSF at $\rho = 160$ μm with $\lambda = 532$ nm. **b**, Simulated images of letter “E” patterns with 40 μm (left) and 20 μm (right) line spacings.

In addition, due to fabrication constraints, the experimentally achievable Q-factor of the thin film is limited. We show that the imaging resolution can be further improved by enhancing the angular selectivity via cascading two tilted thin films. Supplementary Figure 12 compares the two imaging configurations. The first configuration employs a single thin film mounted upright, while the second configuration cascades two identical films, each tilted by the same angle ψ about orthogonal axes.

In both cases, the illumination wavelength is chosen such that the resonance angle corresponds to the direction of direct transmission. A translationally-invariant thin film fundamentally responds to the longitudinal wavevector component $k_z = k_0 \cos(\theta - \psi)$, with its resonance linewidth Δk_z determined by the cavity Q-factor. Here, θ denotes the angle of incidence measured with respect to the optical axis, and k_0 is the free-space wavevector. A small angular deviation $\Delta\theta$ around the resonance shifts k_z , yielding an angular linewidth for a film tilted by an angle ψ as,

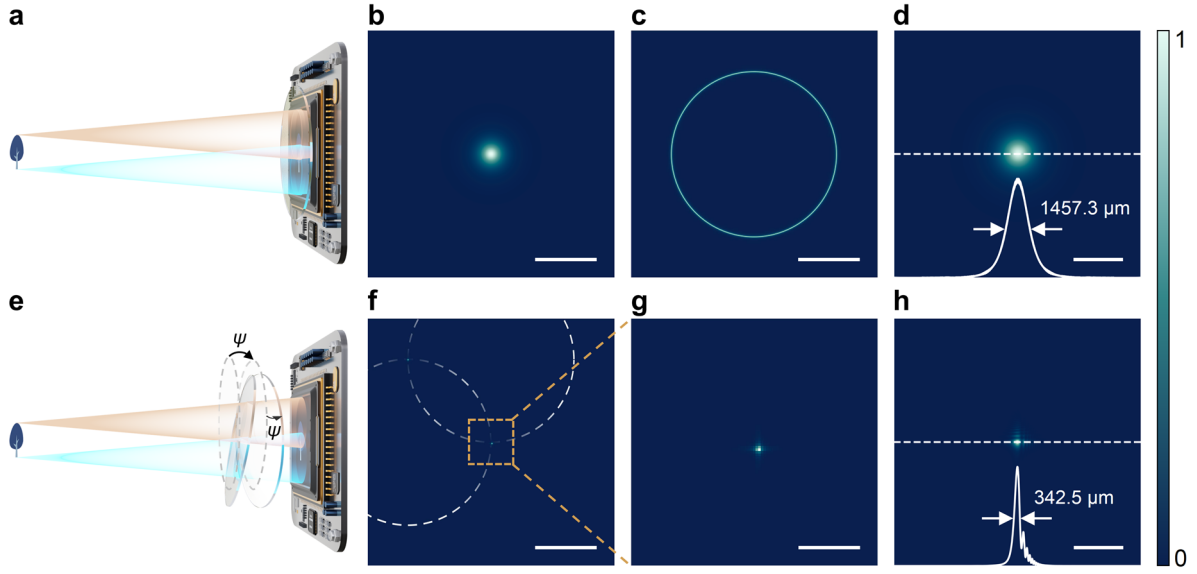
$$\Delta\theta = \frac{\Delta k_z}{k_0 \sin\psi}. \quad (\text{S21})$$

Because $\sin\psi$ grows with increasing ψ , a fixed Δk_z (fixed Q) maps into a smaller $\Delta\theta$ as ψ increases. On the other hand, a single tilted film can achieve a narrow angular linewidth but only along one transverse momentum component, while imaging applications require filtering in both transverse directions. Therefore, two films tilted by ψ about orthogonal axes are cascaded with,

$$t_{\text{total}}(\mathbf{k}) = t(\mathbf{R}_y(\psi)\mathbf{k}) \times t(\mathbf{R}_x(\psi)\mathbf{k}), \quad (\text{S22})$$

where $\mathbf{R}_{\text{axis}}$ denotes the rotation matrix that transforms a vector's coordinates for a rotation of the coordinate system about the specified axis. $\mathbf{k}$ is the incident wavevector of magnitude k_0 . The product of these transmission functions creates a composite transmission response with improved angular selectivity compared to a single film. Supplementary Figure 12b,c show the normalized squared magnitude of the transmission functions of a single thin film operating in dot-mode and ring-mode, respectively. The angular FWHM of the dot-mode (Supplementary Fig. 12b) is 0.81° , and the ring-mode (Supplementary Fig. 12c) is obtained by detuning the wavelength by 1.3 nm from that of the dot-mode. Supplementary Figure 12d presents the PSF corresponding to the dot-mode at an object-to-image distance of 10 cm, with a FWHM of $1457.3 \mu\text{m}$.

Supplementary Figure 12f shows the transmission function of cascading two tilted thin films, with Supplementary Fig. 12g showing the zoom-in view. The dashed rings in Supplementary Fig. 12f mark the individual transmission peak positions of the two orthogonally tilted films. The orthogonally cascaded films reduce the transmission function angular FWHM to 0.073° , an order-of-magnitude narrower compared to the single-film case operated in dot-mode. Furthermore, the off-axis focal spot produced by the cascaded films can be excluded by confining the FOV or the illumination angle. Consequently, as shown in Supplementary Fig. 12h, the PSF generated by the cascaded films is substantially sharper than the PSF generated by a single film, with a FWHM of $342.5 \mu\text{m}$, significantly improving imaging resolution. Taken together, these results indicate that the resolution reported in the main text corresponds to the specific implementation we adopted in the current prototype, and that the practically achievable resolution can be improved for a given task through appropriate design choices.



Supplementary Figure 12 | Resolution enhancement by cascaded tilted thin films. **a**, Schematic of the single upright thin film configuration. **b**, Normalized squared angle-dependent transmission function of the single upright thin film under dot-mode operation. Scale bar: $0.05k_0$. **c**, Normalized squared angle-dependent transmission function of the single upright thin film under ring-mode operation. Scale bar: $0.05k_0$. **d**, PSF for the single upright film configuration. Scale bar: $3000 \mu\text{m}$. **e**, Schematic of two cascaded thin films tilted by ψ about orthogonal axes. **f**, Normalized squared transmission function of the cascaded films. The dashed rings mark the individual transmission peak positions of the two orthogonally tilted films. Scale bar: $0.05k_0$. **g**, Zoomed-in view of the central region in (f). Scale bar: $0.01k_0$. **h**, PSF for the cascaded configuration, showing enhanced sharpness. Scale bar: $3000 \mu\text{m}$.

Supplementary Note 11. Tailoring field of view and magnification with curved detectors for telescopic and microscopic imaging

As established in the main text, the nonlocal thin film in a planar configuration provides a fixed unity magnification. To expand the FOV for a given sensor footprint and enable tunable magnification, we consider integrating the nonlocal thin film onto a curved detector with a spherical surface of radius R as shown in Supplementary Fig. 13a. In this configuration, the magnification is directly determined by the ratio of the detector radius R to the object's distance D_0 from the sphere's center.

This approach provides an attractive alternative to conventional refractive telescopic systems as shown in Supplementary Fig. 13b. In traditional refractive telescopic imaging systems, the finite pixel size of the imaging sensor necessitates the use of lenses with long focal lengths to achieve sufficient magnification for the image to be properly sampled, resulting in large TTL. Conversely, by integrating the nonlocal thin film onto a curved detector, magnification is determined by the detector's geometry rather than a long propagation distance. This allows the system to achieve the required magnification while maintaining a minimal TTL, effectively decoupling magnification from system length. Additionally, for a given fixed

detector size, adjusting the radius of curvature enables flexible control of the FOV, further enhancing the adaptability of this approach.

To simplify calculations, we employ a locally quasi-conformal coordinate transformation that reformulates an isotropic, spherically stratified permittivity defined in the laboratory coordinate system as an equivalent stack of laterally homogeneous planar layers. To clearly represent the geometry involved, we first introduce a surface-adapted coordinate system (s, ϕ, τ) as shown in Supplementary Fig. 13a, where τ denotes the radial distance and (s, ϕ) represent the geodesic arc lengths measured along two mutually orthogonal normal-section directions on a spherical surface defined by a constant radius τ , passing through a local reference point. The surface-adapted coordinates are related to the Cartesian coordinates by the forward mapping,

$$x = s \operatorname{sinc} \gamma, \quad (\text{S23})$$

$$y = \phi \operatorname{sinc} \gamma, \quad (\text{S24})$$

$$z = \tau \cos \gamma, \quad (\text{S25})$$

$$\gamma = \frac{\sqrt{s^2 + \phi^2}}{\tau}, \quad (\text{S26})$$

$$\operatorname{sinc} \gamma = \frac{\sin \gamma}{\gamma}. \quad (\text{S27})$$

The determinant of the Jacobian $J = \partial(x, y, z)/\partial(s, \phi, \tau)$ is,

$$\det J = \operatorname{sinc} \gamma. \quad (\text{S28})$$

The induced metric tensor $G = J^T J$ in the (s, ϕ, τ) coordinate basis is,

$$G = \begin{bmatrix} \operatorname{sinc}^2 \gamma + \beta \frac{s^2}{\tau^2} & \beta \frac{s\phi}{\tau^2} & -\frac{s}{\tau} \\ \beta \frac{s\phi}{\tau^2} & \operatorname{sinc}^2 \gamma + \beta \frac{\phi^2}{\tau^2} & -\frac{\phi}{\tau} \\ -\frac{s}{\tau} & -\frac{\phi}{\tau} & 1 + \gamma^2 \end{bmatrix}, \quad (\text{S29})$$

$$\beta = \frac{1 - \operatorname{sinc}^2 \gamma}{\gamma^2}. \quad (\text{S30})$$

Since the transmission distribution in the dot-mode is concentrated within a narrow spherical cap, the small-patch limit, where $|s|, |\phi| \ll \tau$ (equivalently $\gamma \rightarrow 0$), is a valid approximation. Under this limit, the off-diagonal terms approach zero, and the tangential metric reduces to a conformally flat form. Consequently, surfaces defined by constant τ locally approach planar geometry, and $\varepsilon(\tau)$ can be represented as an equivalent laterally homogeneous multilayer. Under this small-patch limit, an incident wave transforms into a spherical wave with a curvature radius $R(D_0 - R)/D_0$.

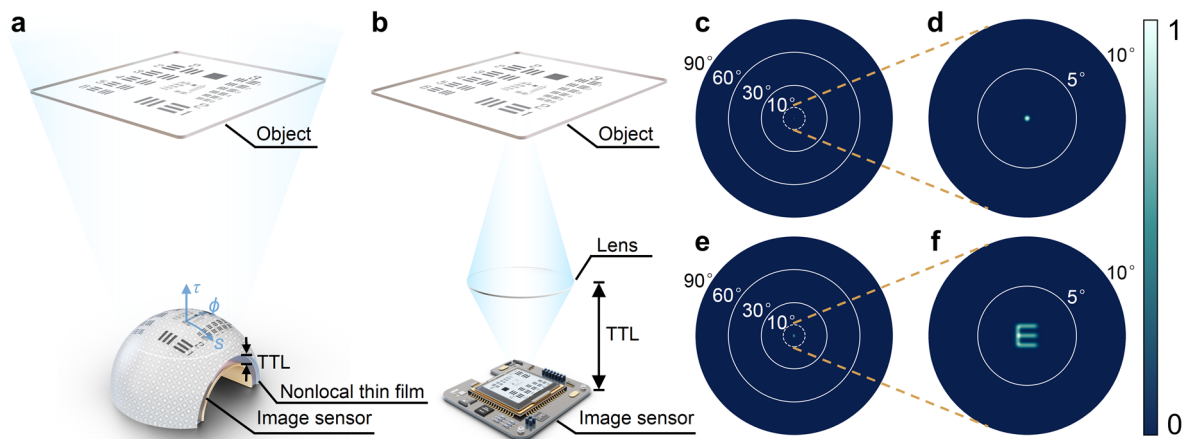
For an originally isotropic medium, the transformed permittivity tensor in the (s, ϕ, τ) coordinate basis is given by the standard transformation-optics relation,

$$\varepsilon_w(s, \phi, \tau) = \varepsilon(\tau) (\det J) G^{-1}. \quad (\text{S31})$$

In the quasi-conformal limit, since the determinant of the Jacobian approaches unity and the metric tensor becomes identity as $(s, \phi) \rightarrow (0, 0)$, the transformed permittivity tensor ε_w remains isotropic and laterally uniform to leading order, with the magnetic permeability μ_w undergoing a similar isotropic transformation within this small-patch approximation.

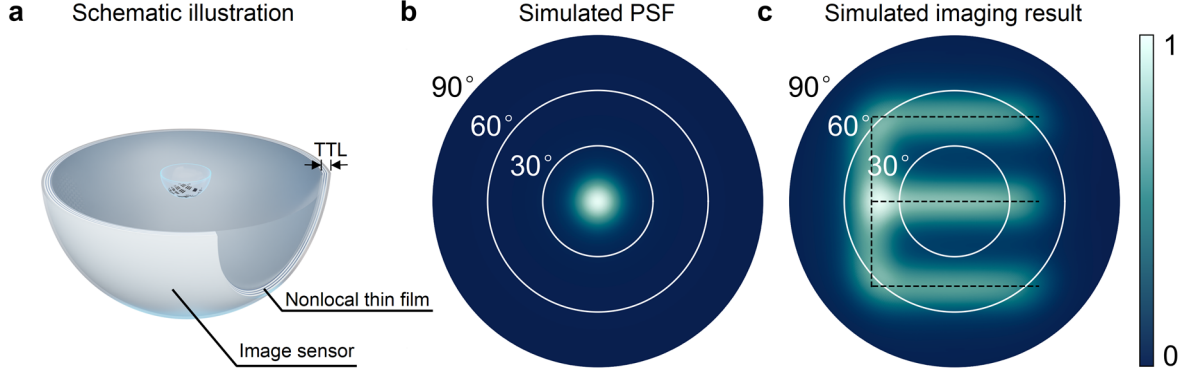
To demonstrate this concept, we simulate a practical implementation where the nonlocal thin film is integrated onto a hemispherical detector with a radius of $R = 5$ cm and the number

of DBR pairs $M = 17$ as shown in Supplementary Fig. 13a. The on-axis PSF in angular coordinates, visualized in Supplementary Fig. 13c with its central region magnified in Supplementary Fig. 13d, exhibits an angular FWHM of 0.48° , corresponding to an object-space resolution of 8.4 cm at a distance of 10 m. The imaging capability is further demonstrated in Supplementary Fig. 13e, which presents a simulated telescopic image of a letter “E” positioned at an object distance of 10 m with a line spacing of 16.9 cm. This configuration yields a magnification of $1/200$. Consequently, the active sensor area of $3.28 \times 2.60 \text{ mm}^2$ for the image sensor in the main text corresponds to an FOV of $656 \times 520 \text{ mm}^2$.



Supplementary Figure 13 | Flexible adjustment of field of view and magnification in curved-detector telescopic imaging without long focal length lenses. **a**, Schematic illustration of curved-detector geometry. **b**, Schematic illustration of conventional telescopic imaging with a long focal length lens. **c**, Polar plot visualization of the on-axis PSF in angular coordinates. **d**, Magnified central region of the PSF from (c). **e**, Simulated telescopic imaging of a letter “E” positioned at an object distance of 10 m with a line spacing of 16.9 cm. **f**, Magnified central region of the simulated “E” from (e).

While the configuration above demonstrates telescopic capabilities, the curved-detector concept can also be adapted for microscopic imaging by utilizing a concave spherical geometry. By integrating the nonlocal thin film onto a concave spherical surface and placing the object near the center of curvature, the system can achieve geometric magnification to improve resolution. To demonstrate this concept, we simulate a case where the nonlocal thin film is integrated onto a hemispherical detector with a radius of $R = 1 \text{ mm}$ and the number of DBR pairs $M = 17$ at an operational wavelength of 532 nm, as shown in Supplementary Fig. 14a. By placing the object on a concentric spherical surface with a radius of $10 \text{ }\mu\text{m}$, the system achieves a geometric magnification of 100. The PSF in angular coordinates, visualized in Supplementary Fig. 14b, corresponds to an object-space resolution of approximately $3.91 \text{ }\mu\text{m}$ characterized by the FWHM. The microscopic imaging capability is further demonstrated in Supplementary Fig. 14c, which presents a simulated image of a letter “E” with a line spacing of $8 \text{ }\mu\text{m}$. This configuration confirms that by leveraging geometric magnification enabled by the curved detector geometry, the system is capable of performing microscopic imaging.



Supplementary Figure 14 | Resolution enhancement in curved-detector microscopic imaging. **a**, Schematic illustration of the curved-detector geometry with a radius of $R = 1$ mm. **b**, Polar plot visualization of the PSF in angular coordinates with DBR pairs $M = 17$. **c**, Simulated microscopic imaging of a letter “E” with a line spacing of $8 \mu\text{m}$ on a concentric spherical object surface of radius $10 \mu\text{m}$.

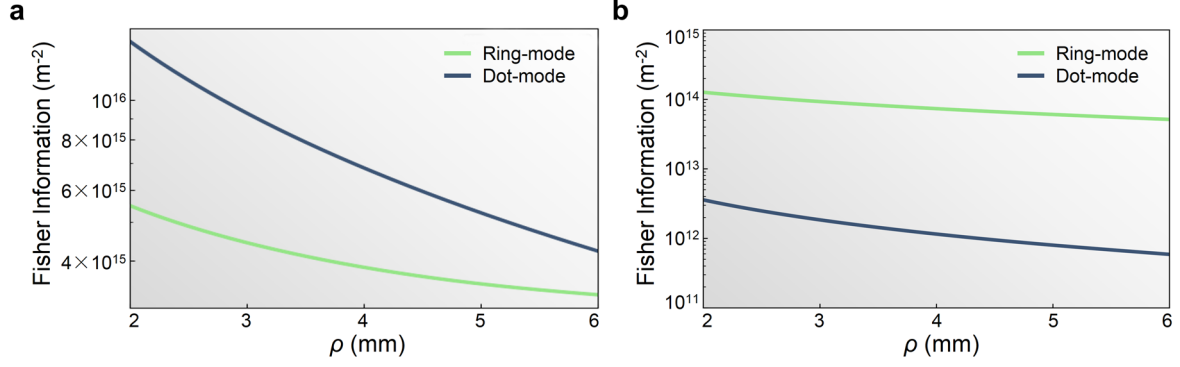
Supplementary Note 12. Fisher-information analysis of dot- and ring-mode PSFs

The Fisher information (FI) sets the Cramér–Rao lower bound (CRLB) on the precision with which the object’s lateral coordinate x and axial coordinate ρ can be inferred from noisy intensity data¹⁸. For a normalized PSF $h(x, y, \rho)$, the FI with respect to a parameter $\zeta \in \{x, \rho\}$ is given by,

$$\text{FI}_{\zeta} = \frac{1}{\sigma_N^2} \sum_i \sum_j \left[\frac{\partial}{\partial \zeta} h(x_i, y_j, \rho) \right]^2, \quad (\text{S32})$$

where σ_N^2 is the noise variance and the summations are taken over the sensor area.

Supplementary Figure 15a, b show the calculated lateral and axial FI for the dot-mode and ring-mode PSFs, respectively, assuming a signal-to-noise ratio of 10 dB. The dot-mode possesses substantially higher lateral FI, confirming its suitability for high-resolution, direct 2D imaging. In contrast, the ring-mode provides significantly greater axial FI throughout the object distance range of 2 mm to 6 mm, indicating superior sensitivity to depth variations. Quantitatively, the CRLB derived from the axial FI yields an axial resolution of $0.09\text{--}0.14 \mu\text{m}$ for the ring-mode PSF over this distance range. These two modes exhibit complementary characteristics, enabling the imaging system to achieve sensitivity to depth variations while simultaneously preserving lateral resolution.



Supplementary Figure 15 | Fisher information analysis for dot- and ring-mode PSFs. **a**, Lateral and **b**, axial Fisher information calculated for both the dot- and ring-mode PSFs, assuming a 10 dB signal-to-noise ratio.

Supplementary Note 13. Detailed method for 3D reconstruction via wavelength switching

Supplementary Figure 16 shows the workflow for physics-informed 3D reconstruction via wavelength switching. First, to reduce background noise, we isolate the region enclosed by the dot-mode image. Specifically, we generate a binary mask Ω by thresholding followed by morphological filtering, which outlines the object's approximate spatial extent. To remain within the PSF calibrated depth range, H is parameterized via a sigmoid. The depth map $H(x, y)$ and albedo $A(x, y)$ are taken as optimization variables. Before computing the soft distribution over depth planes, both H and A are convolved with a Gaussian kernel with a standard deviation of 170.0 μm to suppress high-frequency noise. Each lateral pixel is softly distributed over all depth planes through a temperature-controlled soft-max kernel¹⁹,

$$P_j(x, y) = \frac{\exp\left[\frac{\Gamma_j(x, y)}{\tau}\right]}{\sum_k \exp\left[\frac{\Gamma_k(x, y)}{\tau}\right]}, \quad (\text{S33})$$

$$\Gamma_j(x, y) = |H(x, y) - \rho_j|, \quad (\text{S34})$$

where j indexes the discretized depth layers and the hyper-parameter τ acts as a soft-max temperature that reconciles numerical stability with reconstruction accuracy. P_j denotes the normalized per-depth soft-assignment weight for the plane at depth ρ_j . A larger P_j indicates a higher likelihood that the pixel's depth lies on the plane at depth ρ_j . Γ_j is the pixel-wise absolute depth discrepancy. Multiplying P_j by A yields the grayscale distribution at each depth plane.

Depth-dependent PSF $h_{\text{dot}}(\rho_j)$ and $h_{\text{ring}}(\rho_j)$ are obtained by fitting calibration data acquired before imaging. Convolution every slice with the appropriate PSF and summing over j generates two synthetic sensor images as,

$$I_{\text{dot}}^{\text{sim}} = \sum_j [P_j A] * h_{\text{dot}}(\rho_j), \quad (\text{S35})$$

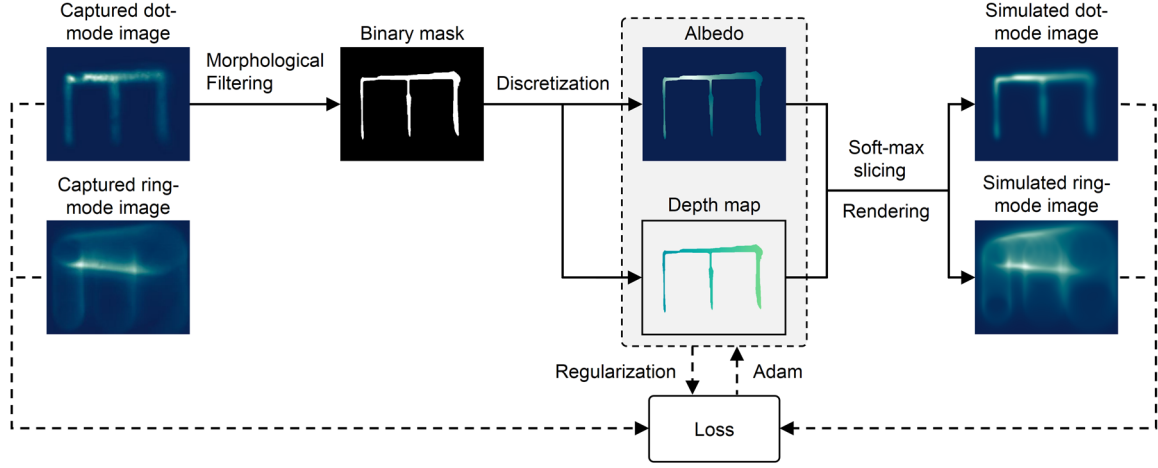
$$I_{\text{ring}}^{\text{sim}} = \sum_j [P_j A] * h_{\text{ring}}(\rho_j), \quad (\text{S36})$$

which are fully differentiable with respect to H and A .

The loss function of the optimization problem is,

$$L(H, A) = \|I_{\text{dot}}^{\text{sim}} - I_{\text{dot}}^{\text{capt}}\|_2^2 + \lambda_r \|I_{\text{ring}}^{\text{sim}} - I_{\text{ring}}^{\text{capt}}\|_2^2 + \lambda_H \|H - \Pi\|_2^2 + \lambda_A \|A\|_{\text{TV}}, \quad (\text{S37})$$

where $\Pi(x, y)$ is a plane fitted to the depth map $H(x, y)$ at each iteration via least squares. The regularization terms constrain the smoothness of the depth map H and the albedo map A . The objective is minimized with the Adam optimizer.



Supplementary Figure 16 | Workflow for physics-informed 3D reconstruction via wavelength switching. A binary mask is derived from the captured dot-mode image by thresholding and morphological filtering to bound the object region. The depth map and albedo are optimization variables used to render simulated dot- and ring-mode images at two wavelengths. The loss combines the squared differences between simulated and captured images with regularization, and all parameters are optimized with Adam.

Supplementary Note 14. Error analysis and mitigation in 3D imaging

In 3D reconstruction, errors mainly arise from temperature-induced resonance wavelength drifts of the thin film, non-uniformity of the film thickness, and truncation of the PSF due to finite sensor size. Temperature fluctuations induce shifts in the central resonance wavelength of the Fabry–Pérot thin film mainly through temperature-dependent changes in refractive index. When the temperature changes to T , the shift in central wavelength is approximated as²⁰,

$$\frac{\Delta\lambda}{\lambda_0} \approx \frac{4}{\lambda_0} \frac{n_{\text{TiO}_2}^T n_{\text{SiO}_2}^T (d_{\text{TiO}_2}^T + d_{\text{SiO}_2}^T)}{n_{\text{TiO}_2}^T + n_{\text{SiO}_2}^T} - 1, \quad (\text{S38})$$

where $n_{\text{TiO}_2}^T$, $n_{\text{SiO}_2}^T$, $d_{\text{TiO}_2}^T$, and $d_{\text{SiO}_2}^T$ denote the refractive indices and thicknesses of the TiO_2 and SiO_2 layers at temperature T , respectively. By inserting typical values^{20, 21}, the temperature dependence of the $\text{SiO}_2/\text{TiO}_2$ film system's central resonant wavelength is approximately - 0.045 nm/°C. Since the thermo-optic response generally exceeds the contribution from thermal expansion, in this approximation only the thermo-optic effect has been considered. Due to the thin film's large surface-area-to-volume ratio, temperature fluctuations are difficult to completely eliminate, resulting in subtle differences between calibrated and measured PSFs due to the film's high Q-factor. Such discrepancies introduce reconstruction errors.

To characterize thermal stability, we numerically evaluated the PSF evolution under temperature-induced resonance drifts. We define the thermal stability window as the temperature range in which the PSF's FWHM changes by less than 10% relative to its value at

the design temperature. For the current SiO₂/TiO₂ film system and an imaging distance of 2.25 mm, this stability window is 4.3 °C. The limited thermal stability window reflects the inherent temperature sensitivity of high-Q microcavities. As a mitigation route, utilizing thin film stacks composed of materials with lower thermal sensitivity can be effective. For example, adopting a Ta₂O₅/SiO₂ film system²⁰ can reduce the temperature-induced shift in central resonance wavelength to approximately 0.012 nm/°C. Consequently, the thermal stability window can be extended to 16.8 °C.

Additionally, the temperature-induced shift also accounts for the slight discrepancy between the operational wavelength for the dot-mode PSF (1607.4 nm) and the center resonance wavelength characterized from the angle-resolved transmittance spectra (1607.7 nm), as the two measurements were performed under different ambient thermal conditions.

Moreover, spatial variations in the thin-film thickness introduce a position-dependent shift in the local resonance wavelength, which manifests as a slight deformation of the PSF across the field. As a result, slight deviations arise between the calibrated and in-situ PSFs, leading to depth-reconstruction errors. Compared to the nanometer-scale uniformity of ion-assisted deposition—constrained by spatial non-uniformities in vapor and ion flux—epitaxial thin film growth enables layer-by-layer growth with sub-monolayer thickness precision. Additionally, residual nonuniformity can be further reduced by performing a spatially varying PSF calibration across the sensor plane, effectively aligning the calibrated PSF with the spatially varying response.

Besides, the maximum ring diameter in ring-mode PSFs, and consequently the Fisher information available for depth estimation, is limited by the sensor's finite active area and the minimum working distance set by the protective cover glass. These factors restrict further improvements in depth accuracy.

Supplementary Note 15. Detailed method for optical diffraction tomography reconstruction via annular illumination

The optical diffraction tomography (ODT) reconstruction was performed using a GPU-accelerated wave-propagation forward model²² that avoids the weak scattering approximation for enhanced accuracy, combined with gradient-based inversion via the Adam optimizer. The thin film transforms the fully incoherent light generated by the rotating diffuser into partially coherent illumination, with the degree of coherence governed by the magnitude squared of wavelength-tuned transmission function $t(\mathbf{k})$. The resulting mutual intensity of the transmitted optical field is given by the inverse Fourier transform as,

$$J(\mathbf{x}_1 - \mathbf{x}_2) = \int |t(\mathbf{k})|^2 e^{i\mathbf{k} \cdot (\mathbf{x}_2 - \mathbf{x}_1)} d\mathbf{k}, \quad (\text{S39})$$

where $\mathbf{k}$ and $\mathbf{x}$ denote transverse spatial wavevector and spatial coordinates, respectively. The nonlocal thin film selectively passes wavevectors corresponding to a specific illumination angle θ . For each tilt angle θ_n , azimuth angles are sampled, and the resulting intensities are summed incoherently to yield the simulated pattern for the corresponding annular illumination. In the forward model, diffraction patterns for each tilt angle ranging from 0° to 20° are computed by propagating the incident field through the estimated refractive index contrast $\Delta n(x, y, z)$.

Starting from an initial guess Δn , the simulated intensities $I_{\text{sim}, n}$ are iteratively matched to the captured patterns $I_{\text{cap}, n}$ by minimizing a weighted mean-squared-error loss,

$$L_{\text{data}} = \sum_n w_n \|I_{\text{cap}, n} - I_{\text{sim}, n}(\Delta n)\|_2^2. \quad (\text{S40})$$

Because a single illumination direction only captures frequency components located on a spherical cap of the Ewald sphere corresponding to first-order scattering, ODT reconstruction inherently suffers from a missing-cone effect. To address the ill-posed nature of the inverse problem, as commonly done in conventional ODT through regularized iterative reconstruction algorithms^{23, 24}, additional regularization terms are included in the loss function. A three-dimensional total variation term is incorporated, which imposes differential smoothness penalties in the lateral and axial directions as,

$$L_{\text{TV}} = \sum \alpha_{xy} (\sqrt{\partial_x \Delta n^2 + \varepsilon_{\text{TV}}^2} + \sqrt{\partial_y \Delta n^2 + \varepsilon_{\text{TV}}^2}) + \alpha_z \sqrt{\partial_z \Delta n^2 + \varepsilon_{\text{TV}}^2}, \quad (\text{S41})$$

where α_{xy} and α_z represent the weights for lateral and axial regularization, and ε_{TV} is a small constant to ensure numerical stability. In addition to smoothness, sparsity in the reconstructed volume is promoted through an ℓ_1 -norm penalty as,

$$L_{\ell_1} = \|\Delta n\|_1. \quad (\text{S42})$$

Furthermore, axial exclusivity constraints are introduced to suppress elongation artifacts by penalizing the spread of refractive index contrast distribution along the depth direction as,

$$L_{\text{excl}} = -\sum_{x, y, z} \left(\frac{\Delta n(x, y, z)}{\sum_{z'} \Delta n(x, y, z') + \varepsilon_{\text{excl}}} \right)^2, \quad (\text{S43})$$

where $\varepsilon_{\text{excl}}$ is a small constant introduced to prevent numerical instability caused by division by zero in the denominator. After each iteration of gradient descent, the refractive index contrast estimate is projected onto the non-negative orthant to ensure physical feasibility as,

$$\Delta n(x, y, z) = \max\{\Delta n(x, y, z), 0\}. \quad (\text{S44})$$

Consequently, the combined loss function is written as,

$$L_{\text{total}} = L_{\text{data}} + \lambda_{\text{TV}} L_{\text{TV}} + \lambda_{\ell_1} L_{\ell_1} + \lambda_{\text{excl}} L_{\text{excl}}, \quad (\text{S45})$$

where λ_{TV} , λ_{ℓ_1} and λ_{excl} are positive scalar weights controlling the influence of each regularization term.

Supplementary Note 16. Analysis and mitigation of the missing-cone effect in ODT

The finite area of the imaging sensor inherently restricts the maximum achievable illumination angle, consequently limiting the angular range available for ODT reconstruction. Additionally, because the center of the diffraction pattern translates laterally on the sensor as the illumination direction sweeps around the annulus, the diffracted pattern produced by one azimuthal component can overlap with the unscattered background produced by the opposite component. This overlap lowers fringe visibility and reduces contrast. The contrast of diffraction patterns progressively decreases as the illumination angle increases. Consequently, the sensor's limited bit precision (8-bit) becomes a bottleneck in resolving these low-contrast signals, thereby constraining the maximum achievable illumination angle. Furthermore, image clipping caused by the protective spacer thickness restricts the captured diffraction intensity patterns. Collectively, these limitations result in the well-known missing-cone effect, characterized by incomplete frequency coverage along the axial direction of the sample's Fourier spectrum.

Specifically, the missing-cone effect arises from the finite angular coverage in the three-dimensional Fourier domain under the weak-scattering approximation, as illustrated in

Supplementary Fig. 17a,d. Each azimuthal direction of the annular illumination corresponds to an individual Ewald shell, and the effective frequency support is the superposition of these shells. The blue region denotes the frequency support and the gray region indicates the missing cone.

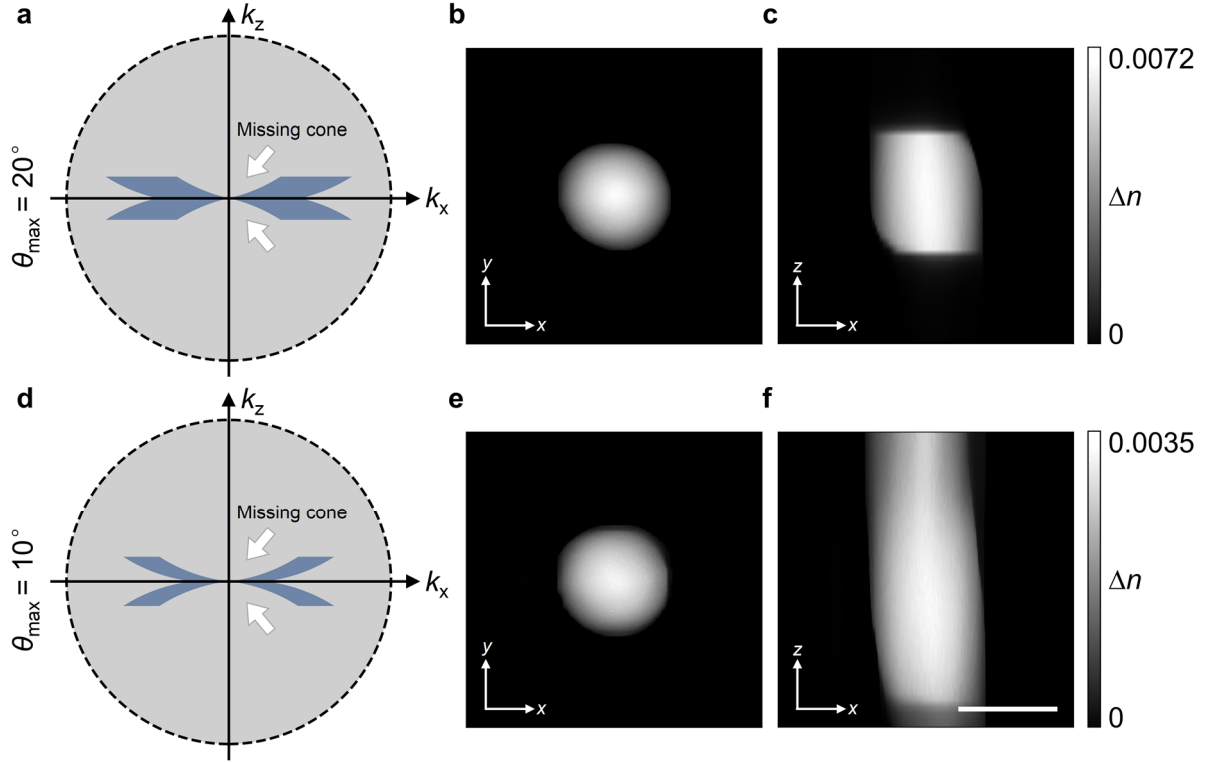
Supplementary Figure 17a-c present the frequency support for refractive index contrast (a), lateral (b), and axial (c) slices through the center of a 200 μm -diameter polystyrene bead reconstructed with a maximum illumination angle of $\theta_{\text{max}} = 20^\circ$ (NA = 0.34), resulting in a measured lateral diameter of 215.3 μm and axial length of 266.6 μm .

In contrast, Supplementary Figure 17d-f depicts analogous results with $\theta_{\text{max}} = 10^\circ$ (NA = 0.17), where the bead exhibits greater axial elongation, showing a lateral diameter of 218.8 μm and axial length of 573.2 μm . These measurements clearly demonstrate that increasing the angular illumination range progressively fills in the missing frequency region, thereby significantly reducing axial elongation artifacts.

Besides these geometric constraints, additional sources of reconstruction errors include thermally induced variations in refractive index and film thickness, which provoke resonance angle drift in the Fabry–Pérot thin film. Such thermal drift disrupts the consistency between the calibrated and actual illumination angles during imaging.

To address these constraints in the current setup, we introduced a circular aperture with a diameter of 1 mm and centered the sample within this opening. By restricting the FOV, the diffracted intensity distribution remains spatially separated from the counterpropagating background, which suppresses background noise and enhances contrast at large illumination angles.

Potential solutions to comprehensively address these limitations and enhance reconstruction accuracy include adopting image sensors with larger active areas and improved bit precision. Directly integrating the thin film onto the CMOS sensor surface or a thinner cover glass can eliminate image clipping caused by spacer thickness, thereby extending usable angular range. Furthermore, incorporating additional imaging wavelengths enriches the frequency information^{25, 26}, effectively reducing artifacts from the missing-cone region. Finally, advanced reconstruction algorithms enhanced by deep learning techniques²⁷ offer a viable approach for compensating incomplete frequency information, further suppressing axial elongation artifacts and improving overall fidelity.



Supplementary Figure 17 | Missing-cone effect and its impact on axial elongation in polystyrene bead reconstructions. **a**, Frequency support for refractive index contrast (blue region) under annular illumination with maximum angle $\theta_{\max} = 20^\circ$ (NA = 0.34), showing the missing-cone region in the gray region. The black dashed circle represents the complete frequency spectrum within the propagating wave range. **b**, **c**, Lateral (**b**) and axial (**c**) slices through the center of a 200 μm -diameter polystyrene bead reconstructed under $\theta_{\max} = 20^\circ$. **d**, Frequency support for refractive index contrast (blue region) under annular illumination with maximum angle $\theta_{\max} = 10^\circ$ (NA = 0.17), showing the missing-cone region in the gray region due to limited angular coverage. The black dashed circle represents the complete frequency spectrum within the propagating wave range. **e**, **f**, Lateral (**e**) and axial (**f**) slices through the center of a 200 μm -diameter polystyrene bead reconstructed under $\theta_{\max} = 10^\circ$. Scale bars: 200 μm .

Supplementary Note 17. Light throughput quantification and efficiency optimization

To quantify the light collection efficiency of the system, we introduce the effective f -number. In a conventional imaging system, the image irradiance E is governed by the f -number (N_F), which is the standard metric for quantifying the light-gathering capability, following the relation $E \propto L/N_F^2$, where L is the source radiance²⁸. For our translationally-invariant system, the light collection is governed by the squared angle-dependent transmission function. Unlike conventional lenses which possess a clear physical aperture, the transmission of our film varies continuously with the incident angle. To enable a direct comparison using the standard f -number metric, we normalize this continuous response to an equivalent top-hat function. We define an effective angular width $\Delta\theta_{\text{eff}}$ as the diameter of an equivalent top-hat angular response that yields the same total integrated energy as the transmission profile. The resulting image

irradiance is given by,

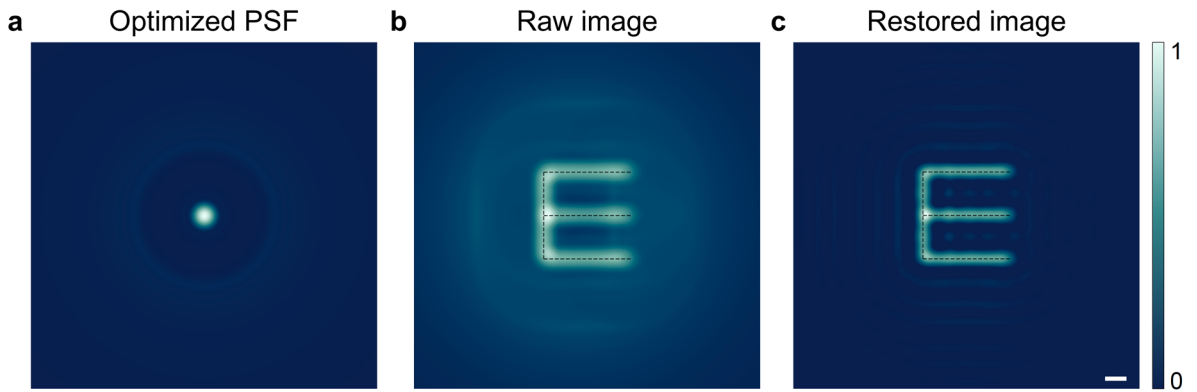
$$E = L\pi\sin^2(\Delta\theta_{\text{eff}}/2). \quad (\text{S46})$$

By equating these two expressions, the effective f -number is determined as,

$$N_F = \frac{1}{2\sin(\Delta\theta_{\text{eff}}/2)}. \quad (\text{S47})$$

Substituting $\Delta\theta_{\text{eff}} = 1.22^\circ$ for the nonlocal thin film, we obtain $N_F = 47.0$. Compared to a typical telecentric lens with an f -number of 8^{29-31} , the light collection efficiency of our system is approximately 1/35 of that of standard optics. We fully acknowledge that the proposed imaging platform is not a universal replacement for conventional optics in all scenarios. Due to the narrowband/dispersive nature of the system, it may be particularly suited for laser-based imaging applications.

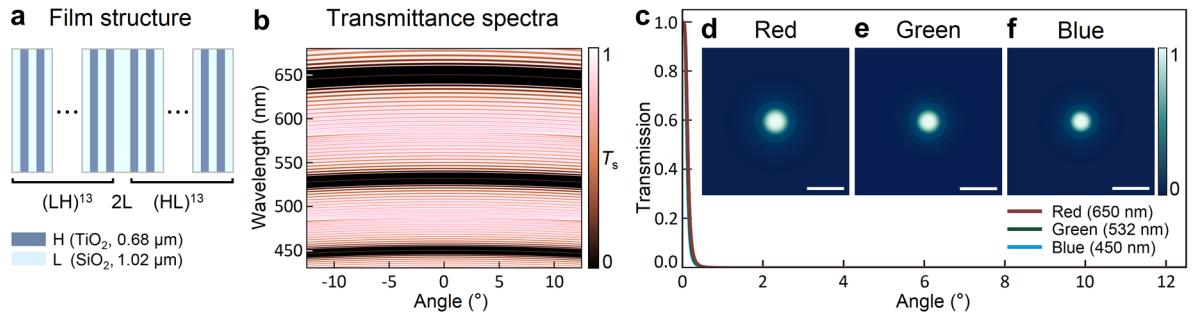
To further improve the efficiency, one potential strategy is to engineer a more complex PSF structure, rather than a single spot, at the cost of requiring deconvolution. Specifically, we optimized the system design by increasing the central layer thickness to $141.5\ \mu\text{m}$ and setting the number of DBR pairs to $M = 7$. At an operating wavelength of $1550\ \text{nm}$, this configuration maintains a central spot FWHM of $87.7\ \mu\text{m}$ while introducing a ring with a FWHM of $119.6\ \mu\text{m}$ (Supplementary Fig. 18a). Simulations indicate that this configuration expands the effective angular width $\Delta\theta_{\text{eff}}$ to 3.43° by collecting light through both the central lobe and the side ring. This enhancement reduces the effective f -number from 47.0 to 16.7. Although this configuration introduces a broader PSF with sidelobes, the spatial information is preserved and can be restored using standard linear deconvolution without requiring complex priors. The resulting image is shown in Supplementary Fig. 18b. Applying Wiener deconvolution to the raw image with 20 dB added Gaussian noise successfully recovers the target structure (letter “E”, line spacing $200\ \mu\text{m}$) (Supplementary Fig. 18c).



Supplementary Figure 18 | Simulation of the efficiency-optimized system configuration with a central layer thickness of $141.5\ \mu\text{m}$ and DBR pairs $M = 7$. **a**, The calculated PSF exhibiting a central spot with a surrounding ring. **b**, The simulated image of the target object (letter “E” with a line spacing of $200\ \mu\text{m}$). **c**, The restored image after applying Wiener deconvolution. Scale bar: $100\ \mu\text{m}$.

Supplementary Note 18. Polychromatic design for multiwavelength operation

While the nonlocal Bragg microcavity is intrinsically narrowband, it can be engineered for polychromatic operation by adjusting the structure. The designed multilayer thin film structure (Supplementary Fig. 19a) supports high-order Fabry–Pérot resonances, as shown in the angle-resolved normalized transmittance spectra T_s (Supplementary Fig. 19b). By increasing the physical thickness of the central SiO₂ layer to 2031 nm, the microcavity supports multiple discrete resonances at normal incidence. This design aligns the transmission peaks to the red (650.0 nm), green (531.8 nm), and blue (450.0 nm) wavelengths (Supplementary Fig. 19c), producing dot-mode PSFs for each color channel (Supplementary Fig. 19d–f). It should be noted that transmission sidebands appear at larger incidence angles around 12.5°. To suppress these unwanted peaks, the angular divergence must be constrained, which can be achieved by engineering the surface bidirectional reflectance distribution function (BRDF) or by angular conditioning in the illumination optics.



Supplementary Figure 19 | Polychromatic thin-film design for RGB imaging. **a**, Schematic of the multilayer thin film structure. **b**, Simulated angle-resolved normalized transmittance spectra T_s . **c**, Simulated angle-dependent transmittance showing aligned normal-incidence transmission peaks at the red (650.0 nm), green (531.8 nm), and blue (450.0 nm) wavelengths. **d–f**, Corresponding simulated dot-mode point spread functions (PSFs) for red, green, and blue illumination, respectively. Scale bar: 100 μm.

Beyond employing sources that match the resonant characteristics of the current device, it is also worth exploring whether broadband operation characterized by spectrally invariant angular filtering is possible. Some recent works on angular filtering have shown that it is possible to achieve broadband operation by engineering the photonic band structure of dielectric stacks^{32,33}. These studies demonstrate that by employing strategies such as stacking multilayer films with varying periodicities, it is possible to maintain angular selectivity at normal incidence across a wide spectral range. Therefore, the relationship between thickness and bandwidth is a rich design space. Crucially, even if broadband operation necessitates a moderate increase in film thickness, the advantage of our translationally-invariant architecture remains: the system thickness is decoupled from the pixel count, thereby preserving the ultracompact form factor unattainable with conventional optics.

References

1. Born, M. & Wolf, E. *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light* (Cambridge University Press, Cambridge, 1999).
2. Campbell, F. W. & Gubisch, R. W. Optical quality of the human eye. *J. Physiol.* **186**, 558–578 (1966).
3. Kingslake, R. *Optics in Photography* (SPIE Optical Engineering Press, Bellingham, 1992).
4. Jacobson, R. E., Ray, S. F., Attridge, G. G. & Axford, N. R. *The Manual of Photography* 9th edn (Focal Press, Oxford, 2000).
5. Smith, W. J. *Modern Optical Engineering: The Design of Optical Systems* 4th edn (McGraw-Hill, New York, 2008).
6. Miller, D. A. B. Why optics needs thickness. *Science* **379**, 41-45 (2023).
7. Duparre, J., Wippermann, F., Dannberg, P. & Brauer, A. Artificial compound eye zoom camera. *Bioinspir. Biomim.* **3**, 046008 (2008).
8. Toyama, M. & Takami, M. Luminous intensity of a gradient-index lens array. *Appl. Opt.* **21**, 1013-1016 (1982).
9. Kiesslich, R., Goetz, M., Hoffman, A. & Galle, P. R. New imaging techniques and opportunities in endoscopy. *Nat. Rev. Gastroenterol. Hepatol* **8**, 547-553 (2011).
10. Yokota, T., Fukuda, K. & Someya, T. Recent Progress of Flexible Image Sensors for Biomedical Applications. *Adv. Mater.* **33**, e2004416 (2021).
11. Dahiya, R. S., Metta, G., Valle, M. & Sandini, G. Tactile Sensing—From Humans to Humanoids. *IEEE Trans. Robot.* **26**, 1-20 (2010).
12. Abad, A. C. & Ranasinghe, A. Visuotactile Sensors With Emphasis on GelSight Sensor: A Review. *IEEE Sens. J.* **20**, 7628-7638 (2020).
13. Chen, A. & Monticone, F. Dielectric Nonlocal Metasurfaces for Fully Solid-State Ultrathin Optical Systems. *ACS Photonics* **8**, 1439-1447 (2021).
14. Boominathan, V., Robinson, J. T., Waller, L. & Veeraraghavan, A. Recent Advances in Lensless Imaging. *Optica* **9**, 1-16 (2022).
15. Levin, A., Fergus, R., Durand, F. & Freeman, W. T. Image and depth from a conventional camera with a coded aperture. *ACM Trans. Graph.* **26**, 70 (2007).
16. Asif, M. S., Ayremlou, A., Sankaranarayanan, A., Veeraraghavan, A. & Baraniuk, R. G. FlatCam: Thin, Lensless Cameras Using Coded Aperture and Computation. *IEEE Trans. Comput. Imaging* **3**, 384-397 (2017).
17. Ma, Z., Shih, K. H., Lopez-Zelaya, C. & Renshaw, C. K. Volumetric imaging efficiency: the fundamental limit to compactness of imaging systems. *Opt. Express* **29**, 3173-3192 (2021).
18. Shechtman, Y., Sahl, S. J., Backer, A. S. & Moerner, W. E. Optimal point spread function design for 3D imaging. *Phys. Rev. Lett.* **113**, 133902 (2014).
19. Bridle, J. Training stochastic model recognition algorithms as networks can lead to maximum mutual information estimation of parameters. *NeurIPS* **2**, 211-217 (1989).
20. Kim, S. H. & Hwangbo, C. K. Derivation of the center-wavelength shift of narrow-bandpass filters under temperature change. *Opt. Express* **12**, 5634-5639 (2004).

21. Gülşen, G. & Naci Inci, M. Thermal optical properties of TiO₂ films. *Opt. Mater.* **18**, 373-381 (2002).
22. Brenner, K. H. & Singer, W. Light propagation through microlenses: a new simulation method. *Appl. Opt.* **32**, 4984-4988 (1993).
23. Lim, J. et al. Comparative study of iterative reconstruction algorithms for missing cone problems in optical diffraction tomography. *Opt. Express* **23**, 16933-16948 (2015).
24. Krauze, W. Optical diffraction tomography with finite object support for the minimization of missing cone artifacts. *Biomed. Opt. Express* **11**, 1919-1926 (2020).
25. Zuo, C., Sun, J., Zhang, J., Hu, Y. & Chen, Q. Lensless phase microscopy and diffraction tomography with multi-angle and multi-wavelength illuminations using a LED matrix. *Opt. Express* **23**, 14314-14328 (2015).
26. Wu, X. et al. Lens-free on-chip 3D microscopy based on wavelength-scanning Fourier ptychographic diffraction tomography. *Light Sci. Appl.* **13**, 237 (2024).
27. Chung, H., Huh, J., Kim, G., Park, Y. K. & Ye, J. C. Missing Cone Artifact Removal in ODT Using Unsupervised Deep Learning in the Projection Domain. *IEEE Trans. Comput. Imaging* **7**, 747-758 (2021).
28. Kingslake, R. The Effective Aperture of a Photographic Objective. *J. Opt. Soc. Am.* **35**, 518-520 (1945).
29. Medley, S. S. et al. Periscope - camera system for visible and infrared imaging diagnostics on TFTR. *Rev. Sci. Instrum.* **56**, 1873-1884 (1985).
30. Zhang, S. et al. 4-domain twisted liquid crystal micropolarizer array for visible linear polarization imaging. *Opt. Express* **29**, 43226-43240 (2021).
31. Liu, C. et al. Correction of a digital micromirror device lithography system for fabrication of a pixelated liquid crystal micropolarizer array. *Opt. Express* **30**, 12014-12025 (2022).
32. Qu, Y. et al. Polarization-Independent Optical Broadband Angular Selectivity. *ACS Photonics* **5**, 4125-4131 (2018).
33. Yin, K. et al. Enabling Manufacturable Optical Broadband Angular-Range Selective Films. *ACS Nano* **15**, 19917-19923 (2021).