OPTICS

Lensless imaging with a programmable Fresnel zone aperture

Xu Zhang[1,2,3]†, Bowen Wang[1,2,3]†, Sheng Li[1,2,3], Kunyao Liang[1,2,3], Haitao Guan[1,2,3], Qian Chen[1,2,3]*, Chao Zuo[1,2,3]*

Optical imaging has long been dominated by traditional lens-based systems that, despite their success, are inherently limited by size, weight, and cost. Lensless imaging seeks to overcome these limitations by replacing lenses with thinner, lighter, and cheaper optical modulators and reconstructing images computationally, while facing trade-offs in image quality, artifacts, and flexibility inherent in traditional static modulation. Here, we propose a lensless imaging method with programmable Fresnel zone aperture (FZA), termed LIP. With a commercial liquid crystal display, we designed an integrated LIP module and demonstrated its capability of high-quality artifact-free reconstruction through dynamic modulation and offset-FZA parallel merging. Compared to static-modulation approaches, LIP achieves a 2.5× resolution enhancement and a 3 decibels improvement in signal-to-noise ratio in "static mode" while maintaining an interaction frame rate of 15 frames per second in "dynamic mode." Experimental results demonstrate LIP's potential as a miniaturized platform for versatile advanced imaging tasks like virtual reality and human-computer interaction.

INTRODUCTION

Lensless imaging, which originates from astronomical x-ray observations where traditional lenses are challenging to manufacture (*1*–*4*), has been extensively used in microscopic biomedical (*5*–*12*) and macroscopic (*13*–*15*) imaging due to its lightweight and cost-effective characteristics. By establishing a nonlinear, controllable, and reversible system transfer function between the incident wavefront and the imaging plane, lensless imaging enables the encoding and decoding between two-dimensional intensity and high-dimensional light-field information (*16*–*18*). In a lensless configuration, the optical lens group, typically dominating the volume in conventional imaging systems, is replaced with amplitude or phase masks with a thickness that constitutes only a small fraction of its former volume (*13*). Once again breaking through the volume boundary of traditional optical imaging systems, lensless imaging drives us to rethink the essence of image formation and explore possibilities beyond point-to-point representation (*19*, *20*).

Lensless imaging systems can be generally classified into two categories, static modulation and dynamic modulation. In the static-modulation configuration, an amplitude or phase mask is typically placed over the sensor as a modulator, and then each point source is coded as a designed pattern rather than a single point (*14*, *17*, *21*). The image captured by the sensor is ideally a linear superposition of patterns generated by each point source, and then the scene information can be reconstructed through direct deconvolution (*20*). However, for nonanalytical static masks, the consistency between the designed pattern and the actual manufactured mask remains uncertain, necessitating high-precision point spread function (PSF) calibration at different depths before experiments. In most cases, lensless imaging involves projecting a high-dimensional light field onto a low-dimensional intensity image. The nonequivalence in information is fundamentally responsible for the ill-posedness of inverse problem (*19*, *22*). In practice, static-modulation lensless reconstruction can achieve a unique solution only with the aid of prior constraints such as Tikhonov regularization and total variation regularization (*13*, *14*), as well as mathematical optimization methods like the alternating direction method of multipliers (*13*–*15*), compressive sensing (*19*, *23*, *24*), and learning-based methods (*25*–*34*). However, because of the inflexibility of mask control, incorrect selection of optimal parameters, and inaccurate calibration of the PSF, static-modulation lensless imaging methods often struggle to achieve high-quality imaging results in more complex scenarios (*13*, *14*).

The emergence of programmable devices, such as the spatial light modulator (SLM) (*35*–*38*) and the reconfigurable scattering mask (*39*), offers a promising and potential dynamic-modulation approach to further enhance the imaging quality of lensless imaging. With dynamic-modulation capability, multiple measurements can be recorded in various ways for high-quality reconstruction (*40*–*43*). The static-modulation mask cannot be changed once designed and manufactured, so that multiple measurements can only be used to reduce noise. In contrast, dynamic-modulation lensless imaging based on the programmable mask can capture multiframe complementary measurement encoded with different patterns, which can not only achieve high resolution and signal-to-noise ratio (SNR) reconstruction but also interpret more scene information (*36*, *44*, *45*). However, previous works with SLMs as programmable masks generally face challenges in low-cost planar integration and high-precision nonanalytical PSF calibration, still relying on the reconstruction framework of traditional static-modulation lensless imaging based on mathematical optimization. Fresnel zone aperture (FZA), inspired by the Fresnel zone plate (FZP) for diffraction-coded imaging (*46*, *47*), offers its analytical physical model and isotropic frequency response for dynamic-modulation lensless imaging.

Similarly to incoherent self-interference holography (*48*), the FZA-based lensless imaging method requires multiple measurements

[1]School of Electronic and Optical Engineering, Nanjing University of Science and Technology, No. 200 Xiaolingwei Street, Nanjing, Jiangsu Province 210094, China. [2]Jiangsu Key Laboratory of Spectral Imaging & Intelligent Sense, Nanjing, Jiangsu Province 210094, China. [3]Smart Computational Imaging Laboratory (SCILab), Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China.
*Corresponding author. Email: chenqian@njust.edu.cn (Q.C.); zuochao@njust.edu.cn (C.Z.)
†These authors contributed equally to this work.

encoded by different FZA patterns to achieve high-quality reconstruction, which brings challenges to its FZA pattern switching with different parameters. As mature display devices (*49*, *50*), liquid crystal displays (LCDs) provide low-cost, compact, and pixel-level modulation capacity for FZA-based lensless imaging. Programmable FZA combines the benefits of both LCD and FZA, offering the potential for dynamic modulation and suppression of twin artifacts for lensless imaging. Combining the analytic forward model of the FZA with the dynamic-modulation capability of an LCD, the programmable FZA offers the potential to avoid complex precision calibration, simplify and robustly implement, and increase the flexibility of lensless imaging systems.

Here, we present lensless imaging with programmable FZA (LIP), a high-resolution, enhanced-SNR, and aliasing-free lensless imaging technique designed to produce high-quality images for compact and cost-effective applications. Under discrete sampling of the programmable mask, optimal parameter matching and offset-FZA parallel merging (OPM) were proposed to avoid aliasing artifacts and synthesize high-quality reconstructed results, respectively. To fully exploit the benefits of the compact lensless volume and the flexible programmable mask, we miniaturized the imaging system into a "LIP module" and incorporated adaptive mode switching ("static mode" and "dynamic mode") to cater to diverse usage scenarios. Compared to the conventional FZA-based methods, the LIP module with OPM enhances resolution by 2.5 times, reaching the pixel size limit of the LCD, improves SNR by 3 dB in static mode, and maintains a real-time reconstruction rate of 15 frames per second (fps) in dynamic mode. Experimental results suggest that LIP presents a promising miniaturized platform for various advanced optical imaging tasks, including virtual reality (VR) and human-computer interaction.

RESULTS

Integrated module for the lensless imaging system with programmable FZA

The schematic of the proposed lensless imaging system with programmable FZA is shown in Fig. 1A. The object $o(x,y)$ is illuminated by incoherent light, and each point $p_{x,y}$ on the object $o(x,y)$ is projected onto the sensor plane with a distinct pattern $t(x,y)$, which corresponds to the magnification of the programmable FZA. Because of the shift-invariant property, the image $g(x,y)$ captured by the sensor is a linear superposition of the projected FZA patterns. As a bridge between incoherent imaging and coherent holography, FZA provides a way to record incoherent point sources in quasi-coherent form (*19*, *48*, *47*). Therefore, the form of $g(x,y)$ is similar to that of the incoherent self-interference hologram.

In the traditional FZA-based method (*47*), the spatial division is used for the four-step phase-shift algorithm to reconstruct an image without twin artifact (Fig. 1B and note S1). However, with programmable FZA, LIP can extend the ability to obtain higher frequency information by displaying different offset FZA. Following offset-FZA computational reconstruction, the proposed OPM method fuses all the collected data

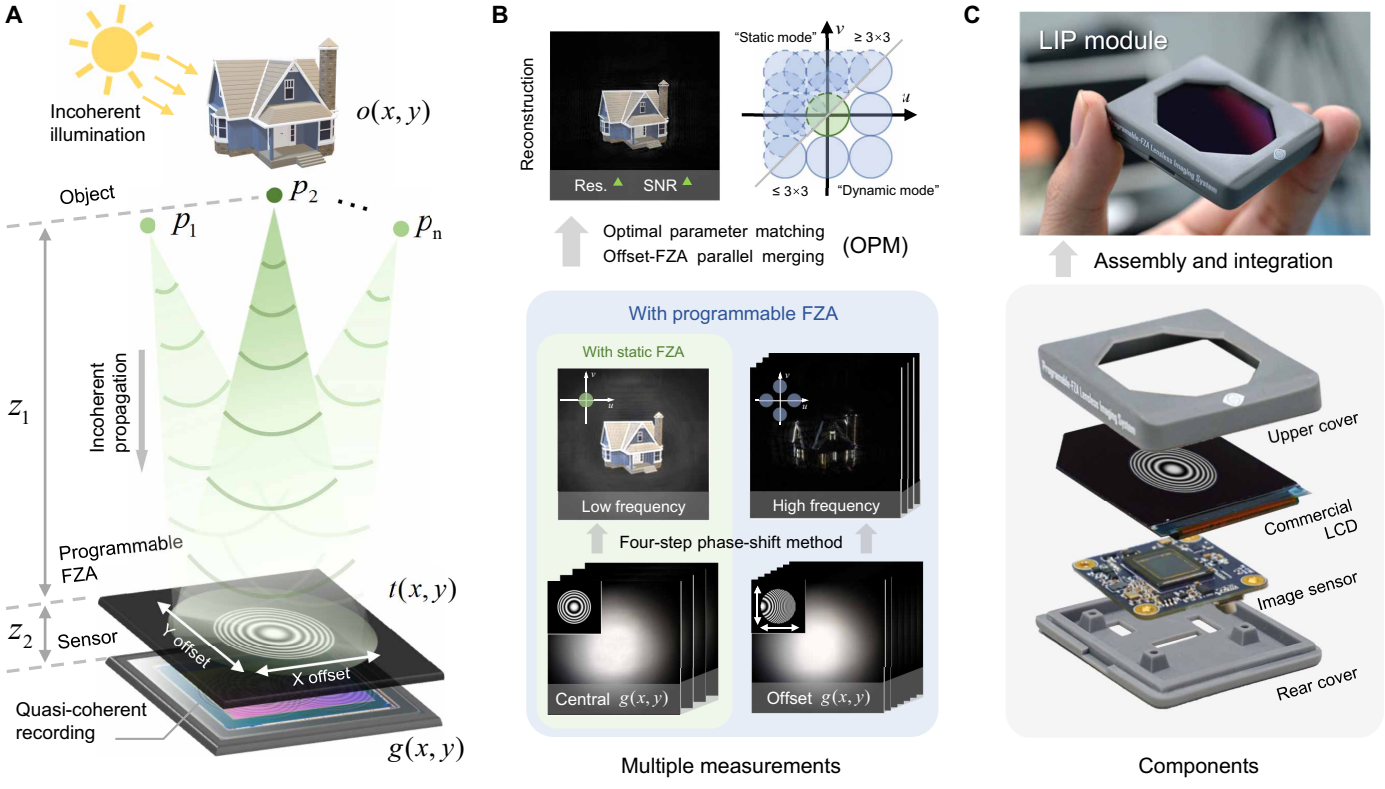

Fig. 1. The forward model of programmable FZA-based lensless imaging system. (**A**) Schematics explaining the basic principle of lensless imaging with programmable FZA. (**B**) Multiple measurements and computational reconstruction with the OPM method, where the designed adaptive mode switching includes static mode and dynamic mode. (**C**) Integrated lensless imaging system LIP module. The system mainly consists of a CMOS image sensor and a commercial LCD, where the LCD can provide programmable amplitude modulation. Photo credit: X.Z.

into one image, resulting in a higher resolution and SNR. In addition, with the flexibility of the programmable mask, we designed adaptive mode switching (Materials and Methods), which can dynamically adjust parameters between high quality (static mode) and high frame rate (dynamic mode) according to different application scenarios, for better adaptability and practicality.

To achieve high-quality lensless imaging with the integrated LIP module, a more detailed forward model has been proposed, considering the programmable FZA's characteristics from the spatial domain and frequency domain perspective (note S2, fig. S1, and movie S1). On the basis of the forward model, we designed an integrated lensless imaging system (Fig. 1C), namely, the LIP module, for experimental validation (Materials and Methods). The LIP module was assembled from a commercial LCD screen as a programmable FZA and a digital camera as an image sensor. All components were assembled in a three-dimensional (3D) printed case, requiring only simple alignment instead of complex and rigorous calibration. Integrating these components eliminates the limitations imposed by bulky lenses, yielding a compact and versatile imaging system.

High-quality lensless imaging with optimal parameter matching and OPM

With the flexibility of programmable FZA, the LIP module has the complete capacity for tuning optical parameters dynamically. However, unlike the previous work, which only used the programmable mask as an ideal continuous mask, we first need to analyze and model its discrete sampling characteristics. The arrangement of pixels in a typical commercial LCD reveals that the actual display area does not attain 100% due to driving circuits (fig. S1B). In particular, as the pixel precision of the LCD increases, the augmentation of the fill factor faces heightened challenges. This, in turn, introduces additional influencing factors into our lensless imaging process, including discrete sampling of patterns and pixel diffraction. Under the discrete pixel sampling of the LCD, the transfer function $T(u, v)$, induced by the pattern $t(x, y)$ displayed on the programmable mask, periodically replicates in the frequency domain with intervals of $[1/\Delta_x, 1/\Delta_y]$, where Δ_x and Δ_y are the pixel pitches in the x and y directions, respectively (Materials and Methods).

The periodic repetition in the frequency domain caused by discrete sampling makes it difficult to further improve the quality of lensless imaging based on programmable masks. The FZA parameter β, which controls the fringe density of FZA, exhibits a linear positive correlation with the imaging resolution in the FZA-based lensless imaging system under a fixed aperture size. As β increases, each replicated spectrum expands until they overlap (fig. S1C). In the spatial domain, this manifests as the PSF's periodic replication and the reconstructed image's mutual overlap. This aliasing problem was not mentioned in previous works, mainly because programmable masks were only exploited as ideal masks, and their frequency response properties were not well modeled before (Materials and Methods).

Here, we propose a method that can use the characteristics of FZA and programmable mask simultaneously to recover the aliased information completely, that is, optimal parameter matching and OPM. OPM is carried out in four steps, as shown in Fig. 2. First, the optimal aperture size R is found according to the maximized angular field of view and the complete pattern sampling, and the maximum optical efficiency is satisfied at the same time (Materials and Methods). Next, we need to choose a β parameter as large as possible without aliasing the spectrum. A larger β will cause the spectrum of periodic replication to be aliased with each other, while a smaller β will not be able to make full use of the frequency domain information, thus reducing the reconstruction efficiency. Only by selecting the optimal β can the periodic replication spectra be exactly independent of each other. Then, optimal parameters are used for subaperture reconstruction without aliasing. In the spatial domain, the FZA is shifted by step length α, corresponding to the spectral sampling position shifted by step length $\beta\alpha/\pi$ in the frequency domain. By designing an appropriate α, it is possible to perform a complete subaperture sampling of the target spectrum. After extracting the spectrum through support constraints, the amplitude and phase information of the subaperture can be recovered without aliasing (for more details about the FZA offset principle, see Materials and Methods). Last, the proposed OPM method synthesizes the information of different frequency components in the frequency domain to reconstruct the complex amplitude information with high quality and high resolution. This spectrum synthesis concept aligns with the ptychography/Fourier ptychography (FP) technique (*51*–*53*). The difference is that FP generally uses the intensity images acquired in the spatial domain as constraints, which requires a high-precision aperture position in the spectrum. Our method combines the phase-shifting method in digital holography to obtain both the intensity and phase information of the subaperture simultaneously, thus directly obtaining the corresponding aperture position in the frequency domain, removing the dependence on the aperture position. After using OPM (note S3 and fig. S2), all the aliasing-free subaperture information is synthesized into one complete spectrum in the frequency domain to realize the high-quality lensless reconstruction without artifacts (for more details about OPM, see movie S2).

Static mode evaluation of the integrated programmable-FZA lensless imaging system

In the static mode, we aim to achieve lensless imaging with superior resolution and higher SNR with the ability to modulate the programmable mask dynamically. However, for a programmable mask, the aliasing problem caused by discrete sampling is the main limiting factor for image quality (see Materials and Methods). We conducted a numerical simulation to explain the connection between the aliasing factors and the FZA parameter β, as shown in Fig. 3A. Here, the aliasing factor is the downsampling scale of the undersampled mask compared to the fully sampled FZA, whose physical meaning is the ratio of the pixel size of the programmable mask to the minimum feature size of the designed FZA mask. It can be observed that when the FZA displayed on the programmable mask avoids aliasing, all reconstructed images are clear and sharp. With an increase in β, the reconstructed spectral range expands, providing more high-frequency information. However, as the aliasing factor and β increase, both the replicated interval Δ_x and the spectral range f_{FZA} expand. This implies that the condition for avoiding aliasing will be violated, causing the periodically repeating FZA spectra to overlap until serious aliasing interference occurs in the reconstructed spatial images (see Materials and Methods). Figure 3B illustrates the reconstructed image quality evaluation index correlation coefficient (CC) and peak SNR (PSNR) under varying β parameters and aliasing factors. It is evident that, compared to scenarios without aliasing, different factors of aliasing have discernible effects on the reconstruction quality. The impact is least pronounced with an aliasing factor of 3, followed by 5. However, at $\beta = 75$ rad/mm^2, the reconstruction quality begins to decline noticeably. With an aliasing factor of 7, the degradation in reconstruction quality

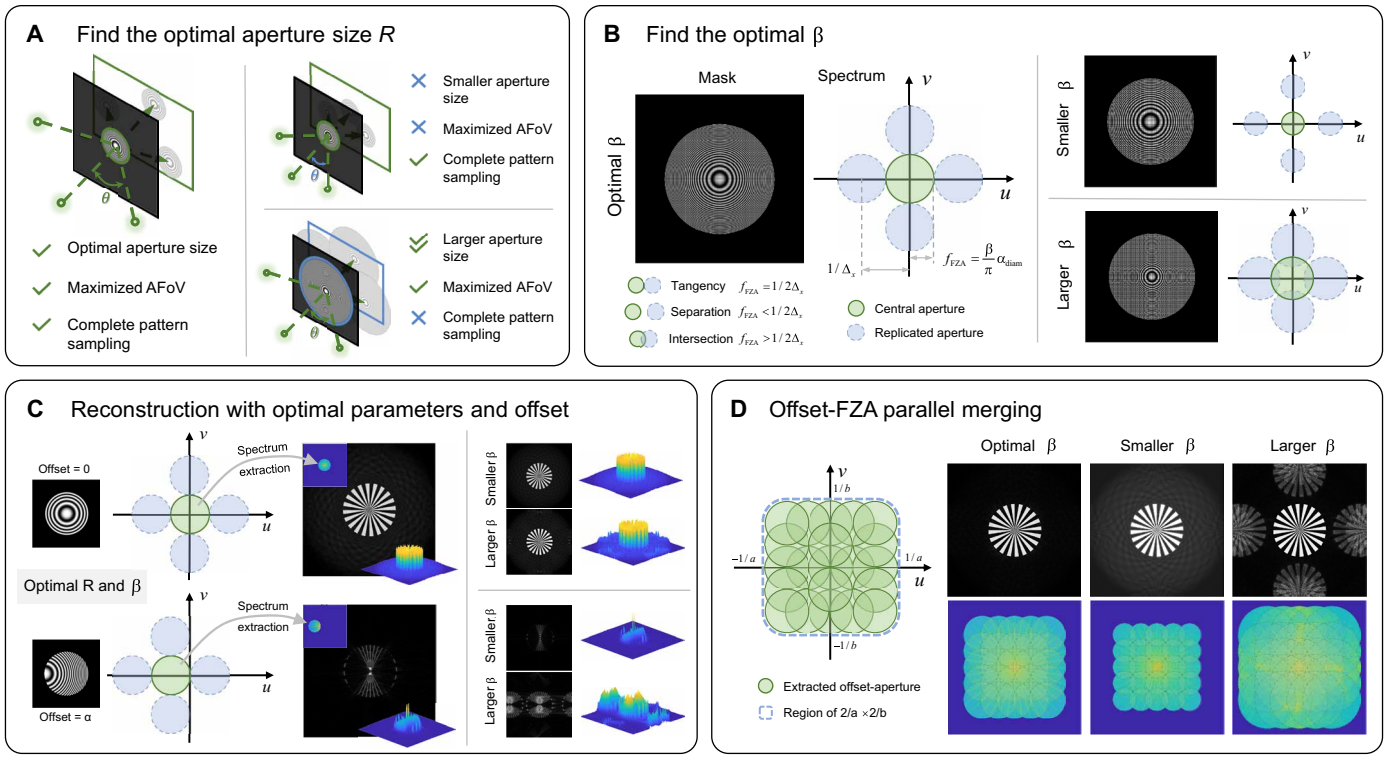

Fig. 2. Optimal parameter matching and offset-FZA parallel merging (OPM) for lensless imaging with programmable FZA. (**A**) Find the optimal aperture size *R*. Maximize the aperture size to achieve the highest light efficiency while ensuring the maximum angular field of view (AFoV) and complete pattern sampling. (**B**) Find the optimal β. Carefully select β to recover the subaperture spectral information without aliasing. (**C**) Reconstruction with optimal parameters and offset. With the optimal parameters, each subaperture's information obtained through the offset FZA can be reconstructed with high quality. (**D**) OPM. Synthesize all subaperture information in the frequency domain to achieve high-quality lensless reconstruction.

starts even earlier, emerging as artifacts as early as β = 30 rad/mm^2. This observation is further corroborated in the frequency domain. Whether it is the reduction in the distance between spectra due to an increased aliasing factor or the broadening of the central spectrum range from an elevated β, the mutual overlap of spectra becomes inevitable, resulting in unavoidable spatial aliasing. This aliasing issue severely constrains the quality of the reconstructed images. The close-up regions of FZA under different pixel sizes (100%, 80%, and 40%) are shown in Fig. 3C. For most programmable masks, their fill factors do not reach 100%, which implies that the cutoff frequency caused by the pixel size mentioned earlier will generally be higher than the replication frequency caused by the pixel pitch. Therefore, even with aliasing present, high-frequency information can still be reconstructed. From the perspective of the PSF, as shown in 3D, under the same β parameter, the reconstruction results with smaller pixel sizes that correspond to narrower PSFs, indicating that despite aliasing, the enhancement of high-frequency details is still achieved, and the true limiting factor for the upper-frequency response is the pixel size rather than the pixel pitch (for more details, see note S4 and fig. S3).

To verify the improvement of the proposed OPM method in terms of resolution, we first simulated the reconstruction of the US Air Force (USAF) resolution target (note S5 and fig. S4). Compared with smaller β and larger β, the optimal β chosen by optimal parameter matching can achieve superior resolution, higher SNR, and narrower PSF with OPM. Besides, more simulations were conducted on 10 groups of complex objects to validate the effectiveness of OPM method, as depicted in Fig. 3E (object "tower," "flower," "bear," and "dog") and fig. S5 (object "loong" to "temple"). Here, the equivalent β is the FZA parameter β required to reach the same spectral range as our proposed method under ideal aliasing-free FZA conditions. The calculation can be expressed as $\beta_{eqv} = N\beta_{opt}$, where N is the spectrum expansion factor corresponding to a spatial-domain resolution enhancement factor. In the simulation, we set the factor of spectrum expansion N to 2.86, i.e., $\beta_{eqv} = 2.86\beta_{opt}$. Comparing the reconstructions under the traditional method with equivalent β, adherence to the optimal parameter matching enabled the extraction of aliasing-free high-frequency information from each subaperture for subsequent reconstruction. The proposed method resulted in a substantial improvement in reconstruction quality, with an average increase of 10.2 dB in PSNR and an average increase of 0.39 in the CC (Fig. 3F). The above numerical simulations show that the proposed forward model of programmable FZA and the corresponding aliasing-free reconstruction method OPM effectively solve the aliasing problem, which is difficult for traditional methods under the programmable mask.

Then, we verified the imaging resolution of the LIP module under the static mode, and the experimental apparatus is illustrated in Fig. 4A. The USAF resolution target was displayed on a monitor 200 mm away from the imaging system, with an actual physical size of 290 mm by 290 mm. Three sets of resolution experiments were performed separately, as shown in Fig. 4B. The optimal β parameter derived from the previously calculated optimal parameter matching (see Materials

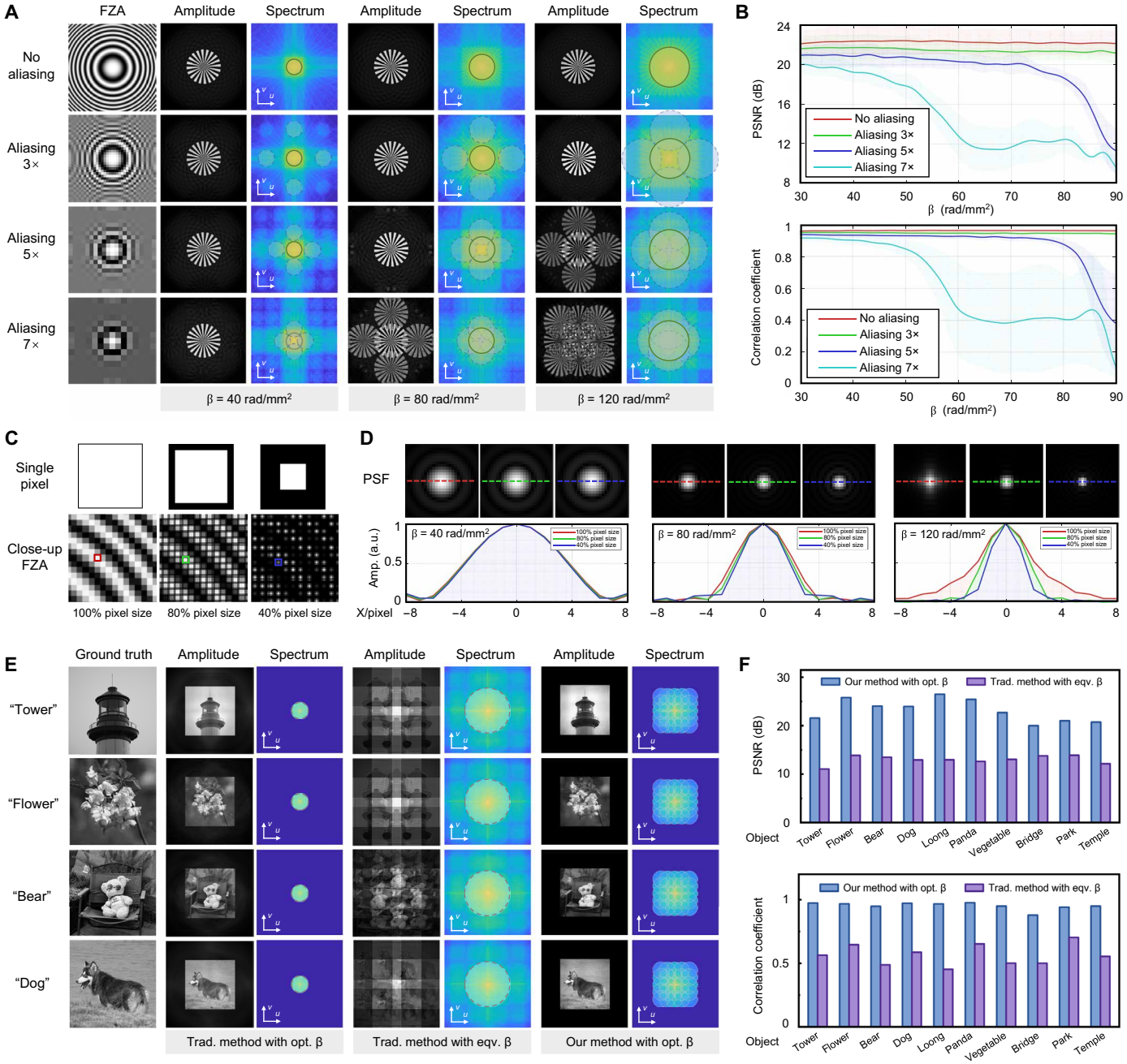

Fig. 3. Numerical simulations on lensless imaging with programmable FZA. (**A**) Reconstruction results with different aliasing factors (no aliasing, aliasing 3×, aliasing 5×, and aliasing 7×) under β = 40, 80, and 120 rad/mm^2. (**B**) PSNR and CC of different aliasing factors' (no aliasing, aliasing 3×, aliasing 5×, and aliasing 7×) reconstruction results under β = 30 to 90 rad/mm^2. (**C**) Close-up FZA under 100, 80, and 40% pixel size. (**D**) Comparison of PSF profiles reconstructed at different LCD pixel sizes (100, 80, and 40%) with β = 40, 80, and 120 rad/mm^2. (**E**) Reconstruction results of the traditional method with optimal β and equivalent β, as well as our method with optimal β for object "tower," "flower," "bear," and "dog." (**F**) PSNR and CC for our method with optimal β and traditional method with equivalent β reconstruction results for object "tower" to "temple." Photo credit: X.Z.

and Methods) was used to image the central aperture using the traditional FZA-based method. It can be seen that the overall imaging quality of the traditional method with optimal β is fair, showing no discernible aliasing. However, the relative resolution exhibited limitations, resolving only up to the line pairs in group −1 element 3, with an SNR of 6.38 dB (for more details about SNR calculation, see note S6). Upon increasing β, an enhancement in resolution was observed, enabling the discrimination of line pairs belonging to group 0 element 4 at equivalent β. Here, the equivalent β is the β parameter required to reach the same spectrum range as our proposed method under ideal aliasing-free FZA. In the experiment, we set the factor of spectrum expansion N to 2.5, i.e., $\beta_{eqv} = 2.5\beta_{opt}$. Nevertheless, the improvement

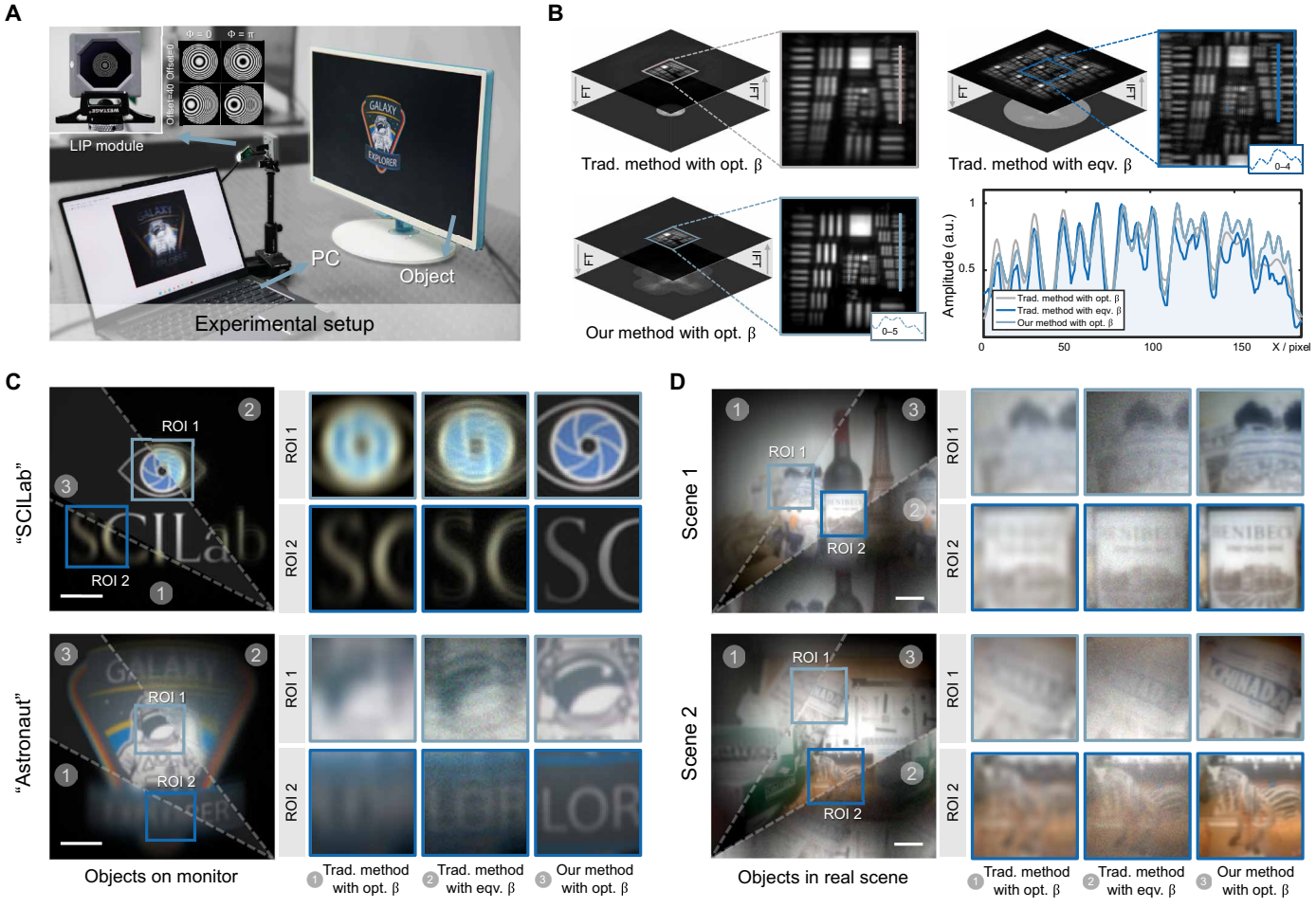

Fig. 4. Performance evaluation of the LIP module in the static mode. (**A**) Experimental setup. The FZA pattern changes during imaging with a specific offset step and phase shift. (**B**) Resolution experiment results using the USAF resolution chart, comparing the traditional method with optimal β and equivalent β, as well as our method with optimal β. (**C**) Reconstruction results of SCILab as a simple color target and Astronaut as a complex color target displayed on the monitor, using the traditional method with optimal β and equivalent β, and our method with optimal β. Scale bars, 50 mm. (**D**) Reconstruction results of two sets of real-world scene experiments, using the traditional method with optimal β and equivalent β, and our method with optimal β. Scale bars, 50 mm.

came at the expense of a considerable decline in overall imaging quality with only 3.63 dB in SNR, marked by artifacts and aliasing infiltrating the central region of the reconstructed image. Subsequently, by introducing the FZA offset with a step size of 40 pixels based on optimal parameter matching, combined with the OPM method using a 9 × 9 aperture, imaging quality that surpasses the equivalent β is ultimately achieved. The improvement is evident in the smoother and more distinguishable line-pair profiles of group −1. In addition, our approach guarantees reconstruction resolution, allowing for the differentiation of group 0 element 5 pairs that are challenging to distinguish using equivalent β and traditional methods, due to the high SNR of 6.65 dB achieved in the reconstruction results (for more details about imaging resolution analysis, see Materials and Methods).

In addition, we conducted experiments involving complex color objects, including "SCILab" as a simple color target and "Astronaut" as a complex color target displayed on the monitor, using the same configuration as the resolution experiments and reducing the number of apertures to 5 × 5 for acquisition efficiency. The reconstruction results and the corresponding close-up regions of interest (ROIs) are presented in Fig. 4C. For the simple target SCILab, compared with the reconstruction results of the traditional method under the optimal β and equivalent β, the proposed method can completely preserve the edge details of the logo (ROI 1) and the text (ROI 2),while greatly reducing artifacts and background noise. Similarly, for the complex target Astronaut, we can also reconstruct the texture of the astronaut's helmet (ROI 1) and the bottom text (ROI 2), which are difficult to recognize with the traditional method.

To further demonstrate the advantages of our method in high-quality lensless reconstruction, we conducted two sets of real-world scene experiments (Fig. 4D). In scene 1, we selected the hat of a snowman doll and the label on a wine bottle as ROIs. As can be seen, in traditional reconstruction methods, the β parameter determined by optimal parameter matching can maximize the reconstruction SNR, but the resolution is still limited. Under the equivalent β parameter, the resolution of the traditional method improves to some extent, but at the cost of notably deteriorating SNR, making previously distinguishable details indiscernible. Our method, while ensuring imaging SNR, also enhances the resolution to the physical limit of the LCD. The details of the snowman doll's hat (ROI 1) and the text "HENIBECK" on the wine bottle (ROI 2) are both visible.

The same conclusion can be verified in the more complex scene 2. The text "CHINADA" in the newspaper (ROI 1) is sharper and clearer than the results of the traditional reconstruction. The stripes on the zebra (ROI 2) have better resolution and SNR under our reconstruction method. It is worth noting that due to the paraxial model approximation and the intensity attenuation of the LCD screen under a large viewing angle, there is some vignetting at the edges of the reconstruction results. The vignetting artifacts can be effectively mitigated through advanced postprocessing correction algorithms (*54*, *55*) or performing intensity compensation based on precalibrated response curves. In brief, the LIP module with the proposed OPM method not only preserves the advantage of a higher SNR at optimal β but also effectively mitigates aliasing artifacts present in reconstruction results under equivalent β, thereby achieving high-quality artifact-free lensless imaging.

Dynamic mode experiment and hand gesture sensing application of the integrated programmable-FZA lensless imaging system

The dynamic mode of the LIP module, unlike the static mode that only considers resolution, requires a more intricate balance between image quality and speed (see Materials and Methods). After considering the response time of the LCD, the speed of sensor acquisition, and the reconstruction time, we have set the imaging rate for the experiments of the dynamic mode at 15 fps with 3×3 apertures. We first built a dynamic scenario consisting of a playing card positioned closer and a hand repeatedly opening and closing at a slightly farther distance, and the results are depicted in fig. S6. A total of 64 frames were captured and reconstructed (fig. S6A), and frame 29 was specifically chosen for the refocusing test (fig. S6B). At the −0.300 m position, neither the playing card nor the hand is on the focal plane. The negative distance here represents a digital refocus from the opposite direction of the incident light. However, at positions −0.334 m and −0.339 m, clear images of the hand and the playing card are obtained, respectively. The dynamic representation of the hand is provided in fig. S6C, showcasing the transition from blurred to clear images as the refocusing position changes (for more details, see movie S3). Compared with the focusing mechanism in traditional lens groups, the refocusing characteristics of the LIP module also pave the way for further reducing the volume of the imaging system and promoting miniaturization and integration application.

Furthermore, the powerful imaging ability and miniaturized size of our LIP module also shed light on volume-limited devices such as VR headsets. As an example, we demonstrated the application of the LIP module in hand gesture sensing and the potential for future integration into VR headsets (Fig. 5). Gesture sensors, being one of the most crucial components in VR devices, offer an intuitive means of interacting with the virtual world and liberate users from the constraints imposed by traditional control joysticks. As shown in Fig. 5A, the gesture sensor is generally located right in front of the VR headset to capture the user's gesture interaction. However, the utilization of a traditional camera module as a gesture sensor often leads to substantial increases in the volume and weight of the VR device, which inevitably compromises the user experience. The traditional camera modules before optimization generally have lens structures more than 20 mm above the sensor surface, while the LIP module's LCD-based programmable mask requires only 1.66 mm, which greatly compresses the imaging system by more than 90%. This size reduction is critical for space-constrained VR headsets, freeing up space for other more important components (Fig. 5B).

To verify the hand gesture sensing capability of the LIP module, we consulted the Apple Inc. developer guidelines and designed four groups of common interactive gestures in VR devices (Fig. 5C), namely, "Tap," "Drag," "Zoom," and "Rotate." The Tap gesture requires the forefinger and thumb to touch each other once, as a click-confirm interaction. The Drag gesture involves moving the hand slowly while holding the forefinger and thumb together, as an interaction to drag a virtual object. The Zoom gesture requires the left and right hands to be separated from each other at the same level, as an interaction to zoom in on a virtual object. The Rotate gesture involves rotating the left and right hands around each other, as an interaction to rotate a virtual object. The results of hand gesture collection and reconstruction are illustrated in Fig. 5D. We collected and reconstructed 16 frames for each gesture and displayed 5 frames among them. Despite the 90% reduction in imaging space, our integrated LIP module can still accurately image and capture hand gestures and ensure that the next step of hand key point detection can be carried out smoothly.

Last, we constructed a gesture interaction scenario to demonstrate the potential of our LIP module further, as shown in Fig. 5E. The user's interaction gestures are captured by the LIP module in front of the user, sent to a PC for recognition, and applied to the virtual object "Earth." As shown in Fig. 5F, the four gestures we designed are well recognized by the hand key-point detection algorithm, allowing the Earth to be selected, dragged, enlarged, and rotated accordingly (for more details, see movie S4). With its adaptive mode switching capability, our LIP module can be integrated into a broader range of sensors on VR headsets, enhancing the user experience with its lightweight and portable design.

DISCUSSION

By incorporating dynamic-modulation capability and the OPM method into the lensless imaging system, we not only address the ill-conditioned problems and twin artifacts associated with static FZA but also mitigate aliasing artifacts caused by programmable masks in the reconstruction process. The high degree of freedom in modulation and flexibility in reconstruction provided by our LIP module allows for adaptive mode switching and parameter tuning tailored to specific application requirements, ensuring optimal encoding and decoding. Last, we achieved a 2.5× resolution improvement (to the pixel size limit of the LCD), 3 dB SNR improvement (static mode), and 15 fps imaging speed (dynamic mode) on our miniaturized system LIP module. Moreover, we demonstrate the potential of the LIP module for hand gesture sensing in volume-limited devices such as VR headsets. Given the widespread adoption of LCD technology in smart devices, our approach provides low-cost, compact, and lightweight imaging capabilities beyond conventional display applications. In essence, the impressive performance and adaptive mode switching of LIP can primarily be attributed to the programmable mask's ability, which efficiently and rapidly maps high-dimensional object information onto collected two-dimensional intensity data in a controlled way. This feature ensures stable, fast, and well-conditioned solutions for the more inverse problem.

However, it must be acknowledged that the primary source of the aliasing artifacts and other issues originates from defects introduced in the LCD with nonideal pixel size and insufficient fill rate. Fortunately,

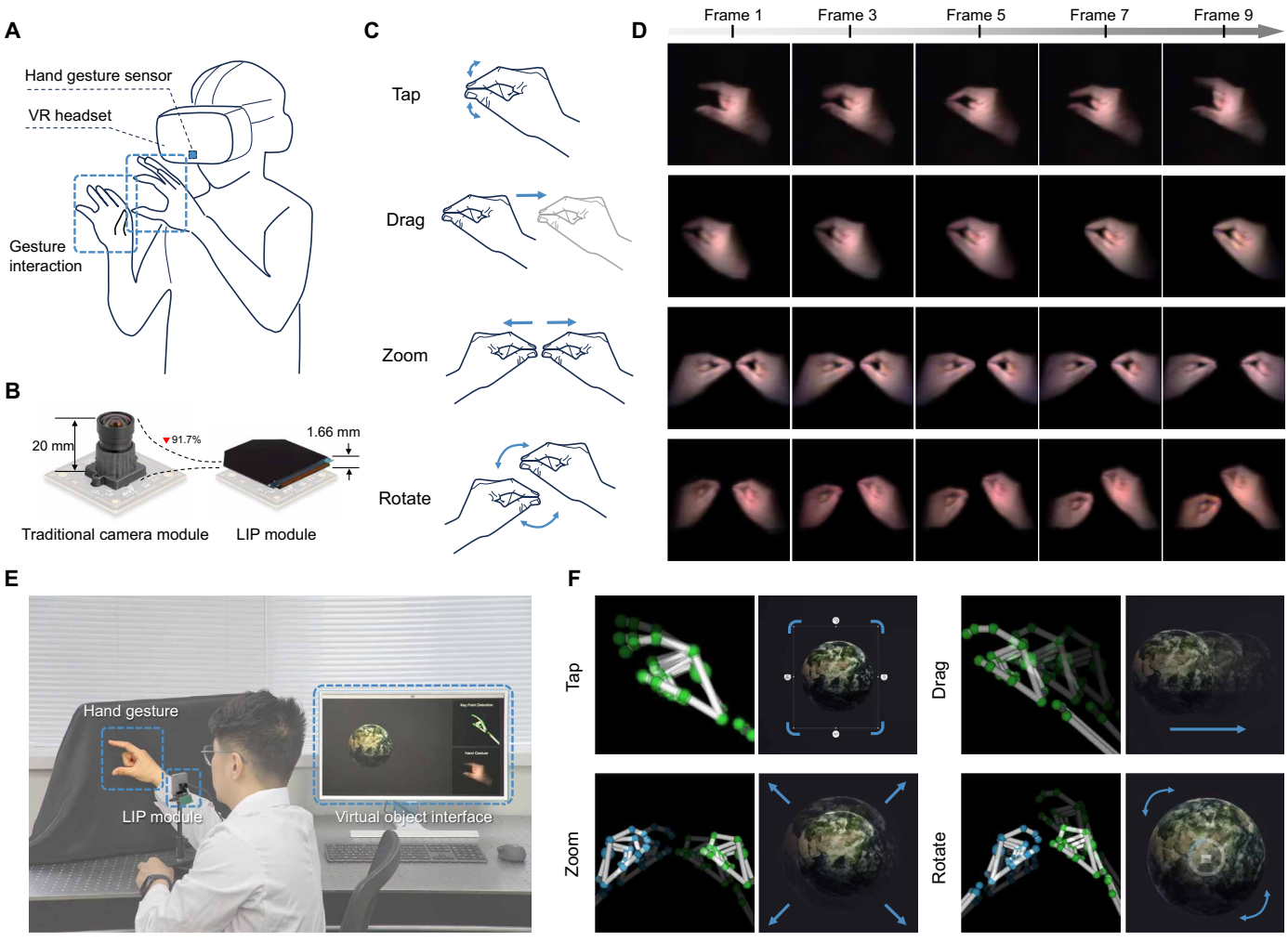

Fig. 5. Application of LIP module with the dynamic mode in hand gesture sensing. (**A**) User gesture interaction scene when wearing a VR headset. (**B**) Thickness comparison between the traditional camera module and LIP module. (**C**) Four groups of common interactive gestures in VR devices, namely, Tap, Drag, Zoom, and Rotate. (**D**) The results of different gesture collections and reconstructions. (**E**) Gesture interaction scenario with LIP module. (**F**) The recognition results corresponding to the four gesture interactions and the virtual object interaction demonstration. Photo credit: X.Z.

the recent rapid development of liquid crystal technology in the field of precise display and optical communication, which brings higher resolution, smaller size, and faster response speed, can further enhance the dynamic-modulation capability of LIP technology and gradually mitigate artifact problems, thereby advancing LIP toward a more practical future. Thanks to the dynamic tuning capability of the programmable LCD and the effective implementation of the OPM method on discrete sampling patterns, the architecture of our proposed LIP module can be easily extended to lensless imaging systems based on different mask configurations (*56*, *57*). By adjusting the mask position, multiview and high-precision depth information can be extracted (*36*). When the mask is adjusted to specific pinhole shapes, computational spectral reconstruction becomes possible (*58*). Moreover, the principles of the OPM method can be extended to other lensless imaging techniques, enabling the acquisition of high-dimensional information across different frequencies and modalities by manipulating various mask configurations.

In the context of technological development for hand gesture interaction in VR headsets, improvements can be sought from two approaches: enhancing performance limits or optimizing other aspects while meeting application requirements. Our LIP module for hand gesture interaction in VR follows the latter approach, achieving a 91.7% reduction in the imaging system's size and improving the user experience in VR headsets while maintaining high-quality gesture recognition. Furthermore, it is important to highlight that current VR headsets typically incorporate infrared illumination to ensure stable lighting conditions and increase the success rate of gesture recognition. The hand gesture interaction experiments presented in our work were conducted under well-controlled lighting conditions, closely simulating real-world usage scenarios. In future work, we plan to integrate compact illumination sources within the LIP module and leverage the programmable LCD to precisely control the light source, enhancing adaptability to varying environments.

In summary, the potential of LIP lies not only in its compact size and cost-effectiveness but also in its capability to capture high-dimensional

light fields in various forms within a two-dimensional intensity plane. This enables the retrieval of high-quality and valuable high-dimensional information. In the future, our objective is to progressively extend the framework of LIP toward miniaturized and multimode imaging systems, facilitating diverse plug-and-play applications ranging from biomedical imaging to human-computer interaction.

MATERIALS AND METHODS

LIP module configuration

The LIP module was assembled from a commercial LCD screen as a programmable FZA and a digital camera as an image sensor (Fig. 1C). The LCD screen used is a BOE VS021XRM-NW0-DKP0, with a pixel resolution of 1600 × 1600, a pixel pitch of 24 μm, and a viewing angle (θ_x, θ_y) of 40°. The image sensor used is a Sony IMX183 with a pixel resolution of 5544 × 3694 and a pixel pitch of 2.4 μm. Considering the optical efficiency and the pixel size of LCD, the actual image acquisition is set to "pixel_binning = 3" (equivalent pixel resolution 1824 × 1216 and equivalent pixel pitch 7.2 μm). To reduce the size of the LIP module, we positioned the LCD screen close to the image sensor, resulting in a mask-sensor distance z_2 of 1.66 mm, equivalent to the thickness of the LCD screen. Following the optimal parameter matching, we computed the optimal aperture size R_m and FZA parameters β_m for the LIP module: $R_m = \left[W/2 - z_2\tan(\theta_y)\right]/m = 2.72$ mm, $\beta_m = \pi m/(2\Delta_y R_m) = 24.24$ rad/mm^2, where $W = 1824 \times 7.2$ μm = 8.86 mm represents the width of the sensor and magnification factor $m = 1.11$ (for more details, see the "Optimal parameter matching for LIP module" section). Further discussion of the sensor measurement overflow and color-space characteristics (*59*) of the raw data is given in note S7 and fig. S7.

In the numerical simulation, we used high-resolution images (1024 × 1024 pixels) as the ground truth to validate the proposed method. The programmable mask had a pixel pitch (Δ_x, Δ_y) of 24 μm, a pixel size (a, b) of 19.2 μm, and a viewing angle (θ_x, θ_y) of 40°. The sensor size was designed to be 7.37 mm by 4.92 mm (pixel pitch 4.8 μm). The object-mask and mask-sensor distances z_1 and z_2 were set at 100 and 1 mm, respectively, resulting in a magnification factor $m = 1.01$. A discussion of programmable LCD transmittance is given in note S8 and fig. S8.

Adaptive mode switching between the static mode and the dynamic mode

We designed an adaptive mode switching for the LIP module, which is, respectively, the static mode and the dynamic mode. In the dynamic mode, by selecting the optimal parameters β_m and R_m through the optimal parameter matching, the LIP module can gather information from the nine subapertures in the multiplexing measurement. Here, unlike the simulation, where each aperture is individually captured, in the actual experiment, we took into account the limited magnification resulting from the much shorter distance between the mask and the sensor compared to the distance from the object to the mask. As a result, the reconstructed image occupies only about one-ninth of the sensor area. Therefore, we adopt the 3 × 3 spatial multiplexing method (*47*, *60*) to maximize the sensor area, improve the image acquisition efficiency, and adjust the optimal aperture size R_m and FZA parameter β_m to one-third and three times the design parameters, respectively, to ensure the resolution. We captured four different FZA encoded measurements for each frame reconstruction, used the phase shift algorithm for phase retrieval and twin image suppression, and simultaneously resolved the spectral information of nine subapertures. Then, we used the OPM method to perform frequency synthesis without iteration in the frequency domain and, last, executed an inverse Fourier transform back to the spatial domain to obtain the intensity and phase of the object. The LCD used has a refresh rate of 120 Hz, theoretically allowing for high-quality dynamic imaging at frequencies ranging from 1 to 30 Hz. In the actual dynamic scene experiment, the LCD's refresh rate was set to 60 Hz, and the dynamic imaging frame rate was set to 15 Hz to ensure an exposure time of up to 16 ms and acceptable imaging SNR. In the results shown in Fig. 5, we set the sensor exposure to 10 ms and the gain to 6 dB to ensure a certain SNR while reducing motion artifacts. The camera and LCD screen were triggered by a synchronized signal generated by our self-developed circuit.

In the static mode, to further enhance imaging quality, the OPM algorithm incorporates the concept of aperture overlap from FP and resolves the cophasing problem by overlapping the apertures, enabling high-quality lensless reconstruction. Unlike the dynamic mode, which requires only nine apertures, the static mode necessitates 5 × 5 or more apertures for reconstruction, resulting in approximately a triple acquisition time. Thus, the highest imaging rate for the static mode is 5 Hz under the multiplexing method. In the static scene results shown in Fig. 4, to characterize the upper limit of imaging resolution and SNR of the OPM method, we displayed each aperture separately instead of multiplexing nine apertures, and set the sensor exposure settings and gain to 16 ms and 5.5 dB, respectively; other settings were consistent with the dynamic mode.

Discrete sampling of programmable mask

When projecting a designated pattern on a programmable mask, it is imperative to discretize the ideal continuous image and present it in each pixel, ensuring consistent intensity within each pixel region. Given the uniform sampling in both the x and y directions of the LCD, we define the sampling intervals (pixel pitch) in these directions as Δ_x and Δ_y, respectively. Moreover, when considering LCD pixels with a defined area, the scenario becomes more intricate. Assume that each pixel is a rectangle of dimensions $a \times b$ (pixel size). Consequently, the discrete sampling function $s(x, y)$ introduced by the LCD screen can be expressed as follows

$$s(x,y) = \left[\mathrm{rect}\left(\frac{x}{a}\right)\mathrm{rect}\left(\frac{y}{b}\right)\right] * \left[\mathrm{comb}\left(\frac{x}{\Delta_x}\right)\mathrm{comb}\left(\frac{y}{\Delta_y}\right)\right] \quad (1)$$

When the ideal pattern to be displayed is denoted as $t(x, y)$, the actual pattern $t_s(x, y)$ resulting from the discrete sampling on the LCD screen can be expressed as

$$\begin{aligned} t_s(x,y) &= s(x,y)t(x,y) \\ &= \left[\mathrm{rect}\left(\frac{x}{a}\right)\mathrm{rect}\left(\frac{y}{b}\right)\right] * \left[\mathrm{comb}\left(\frac{x}{\Delta_x}\right)\mathrm{comb}\left(\frac{y}{\Delta_y}\right)t(x,y)\right] \end{aligned} \quad (2)$$

Using the convolution theorem of the Fourier transform, the corresponding transfer function $T_s(u, v)$ is

$$\begin{aligned} T_s(u,v) &= F\left\{\left[\mathrm{rect}\left(\frac{x}{a}\right)\mathrm{rect}\left(\frac{y}{b}\right)\right] * \left[\mathrm{comb}\left(\frac{x}{\Delta_x}\right)\mathrm{comb}\left(\frac{y}{\Delta_y}\right)t(x,y)\right]\right\} \\ &= ab\,\mathrm{sinc}(au)\,\mathrm{sinc}(bv)\sum_{m=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} T\left(u-\frac{m}{\Delta_x}, v-\frac{n}{\Delta_y}\right) \end{aligned} \quad (3)$$

This implies that, under the discrete pixel sampling of the LCD, the transfer function $T(u, v)$, induced by the pattern $t(x, y)$, periodically replicates in the frequency domain with intervals of $[1/\Delta_x, 1/\Delta_y]$. When $T(u, v)$ is a band-limited function, and each replicated spectrum remains nonoverlapping, accurate extraction of the complete spectrum can be achieved through band-pass filtering. This allows for precisely restoring the ideal transfer function $T(u, v)$.

Aliasing resulting from the discrete sampling of programmable FZA

Considering the forward model for a programmable mask with an FZA pattern, we can derive the frequency response characteristics of programmable FZA

$$T_{\text{P-FZA}} = ab\,\text{sinc}(au)\,\text{sinc}(bv)\sum_{m=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} T_{\text{FZA}}\left(u-\frac{m}{\Delta_x}, v-\frac{n}{\Delta_y}\right) \tag{4}$$

The aliasing problem arises when the range of the equivalent transfer function $T_{\text{P-FZA}}$ of the FZA is larger than its periodic interval $[1/\Delta_x, 1/\Delta_y]$. When β increases, the frequency range that is repeated at intervals of $[1/\Delta_x, 1/\Delta_y]$ in the frequency domain also increases sequentially, until they overlap. In this process, unwanted frequency components from the surroundings gradually mix into the central spectrum, resulting in disordered phase information for the frequency range corresponding to the overlapping part. In our forward model, the discrete sampling of mask in the spatial domain corresponds to the aliasing in the frequency domain, and the mutual overlap of the spectra will also affect the intensity distribution in the spatial domain, so the PSF will also periodically repeat like the frequency domain. However, it is essential to note that, despite overlap, this does not imply the absence of high-frequency information within the central transfer function. When we further reduce the pixel size of the programmable mask, the cutoff range of the two-dimensional sinc's zero points $2/a \times 2/b$ will expand. This is counterintuitive, as we typically tend to equate pixel pitch with pixel size or solely consider pixel pitch, presuming that reducing pixel pitch is the only method to enhance resolution by recovering high-frequency information before aliasing.

Frequency offset characteristics and spectrum range of FZA

Lensless imaging based on FZA shares fundamental parallels with holography. In holography, the sample details are embedded within the holographic image's high-frequency components or the interference pattern's dense fringes. Likewise, the cutoff frequency of an ideal lensless imaging system is contingent on the density of FZA fringes exhibited within the aperture. The denser the fringes, the better the recorded information for reproducing high-frequency details of the object. The similarity implies that, if we can modulate high-frequency fringes, there is the potential to retrieve intricate details of the object's high-frequency information.

As the FZA undergoes spatial shifting, the density of intensity fringes also varies. During a four-step phase shifting process, the equivalent phase factor similarly shifts (note S2 and fig. S1A). According to the Fourier transform properties of the exponential function, a corresponding shift in the central position occurs in the frequency domain (fig. S1A)

$$T_{\text{FZA}}(u, v) = F\left\{\exp\left[\text{j}\beta\left(x^2+y^2\right)\right]\right\} = \exp\left[j\frac{\pi^2}{\beta}\left(u^2+v^2\right)\right] \tag{5}$$

$$\begin{aligned} T'_{\text{FZA}}(u, v) &= F\left\{\exp\left\{\text{j}\beta\left[\left(x-x_0\right)^2+\left(y-y_0\right)^2\right]\right\}\right\} \\ &= \exp\left[\text{j}\beta\left(x_0^2+y_0^2\right)\right]F\left\{\exp\left[\text{j}\beta\left(x^2+y^2\right)\right]\exp\left[-j2\pi\left(\frac{\beta x_0}{\pi}x+\frac{\beta y_0}{\pi}y\right)\right]\right\} \\ &= \exp\left[\text{j}\beta\left(x_0^2+y_0^2\right)\right]T_{\text{FZA}}\left(u+\frac{\beta x_0}{\pi}, v+\frac{\beta y_0}{\pi}\right) \end{aligned} \tag{6}$$

where T_{FZA} is the equivalent transfer function after the four-step phase shifting, and T_{FZA} is the transfer function after an (x_0, y_0) offset of the FZA. The offset of the original FZA in the spatial domain after the four-step phase shifting can be observed to be equivalent to the phase shift of T_{FZA}, while in the frequency domain, it corresponds to the translation of the transfer function. The spatial and frequency domain offset ratio is $-\beta/\pi$.

After considering the effect of magnification for the actual experimental system due to the distance between the mask and the sensor, we calculated the actual spectral offset size f_{step} of the reconstructed image with different FZA offset sizes and the spectral sampling range f_{diam} corresponding to the aperture diameter as follows

$$f_{\text{step}} = \frac{\beta}{\pi} m\alpha_{\text{offset}} \tag{7}$$

$$f_{\text{diam}} = \frac{\beta}{\pi} m\alpha_{\text{diam}} \tag{8}$$

where α_{offset} and α_{diam} are the offset step and the aperture diameter R on the LCD, respectively. To ensure data redundancy in static mode, we selected a 39.1% subaperture overlap rate as the basis for designing the offset step size, a choice well justified in conventional FP microscopy (FPM) (*61*). At this overlap rate, we can calculate that $\alpha_{\text{offset}} = \alpha_{\text{diam}}/2$, and only the intensity information of subapertures is used for high-precision reconstruction. In the dynamic mode, to improve imaging efficiency, subapertures will no longer overlap; that is, $\alpha_{\text{offset}} = \alpha_{\text{diam}}$. The complete spectrum information of the subaperture can be obtained directly by Fourier transform using the complex amplitude reconstructed by the subaperture, and the reconstruction result can be obtained by simple superposition.

Optimal parameter matching for LIP module

To achieve optimal sampling, we consider the following three aspects in determining the optimal parameters β_m and R_m:

1) Aliasing-free spectrum sampling: The primary cause of aliasing arises from the mismatch between the replication period of frequency domain convolution, dictated by the pixel pitch of the LCD, and the spectral range determined by FZA parameter selection. Therefore, it is crucial to ensure that each replicated spectrum is entirely free from mutual aliasing to optimize the efficiency of subaperture data collection.

2) Maximized angular field of view: The system's angular field of view is primarily determined by the LCD's inherent viewing angle characteristics. The viewing angle limits the maximum angle of light the imaging system can receive, with light intensity diminishing as the angle increases. When selecting the optimal parameters, we maximize the utilization of the LCD's viewing angle to achieve the widest possible angular field of view for the system.

3) Complete pattern recording: If the pattern of the object projected on the sensor through the LCD is not adequately recorded,

the reconstruction quality will degrade radially from the center to the periphery, resulting in a ringing effect at the edges. The issue primarily stems from the improper selection of the FZA aperture size. Hence, it necessitates a trade-off between complete pattern recording and maximizing light efficiency.

Assume the LCD has x and y directional viewing angles of θ_x and θ_y, respectively, and the sensor dimensions are $L \times W$. To satisfy the second and third conditions, the pattern generated at the maximum incident angle must fall within the sensor's range, ensuring complete sampling by the sensor (Fig. 2A). The requirement can be expressed by the following equation

$$R \le \min\left\{\left[L/2 - z_2\tan(\theta_x)\right]/m, \left[W/2 - z_2\tan(\theta_y)\right]/\mathrm{m}\right\} \quad (9)$$

To maximize the light utilization rate, R is generally the maximum value, and for a general sensor, $W \le L$, that is

$$R_m = \left[W/2 - z_2\tan(\theta_y)\right]/\mathrm{m} \quad (10)$$

Next, we consider aliasing free sampling of the spectrum. According to the analysis of FZA, the diameter of the spectral range of FZA is $f_{\mathrm{FZA}} = 2\beta R_m/\pi$. To achieve aliasing-free sampling, there must be no overlap between the spectral ranges of the periodically recurring FZA

$$f_{\mathrm{FZA}} \le \min\left\{\frac{1}{\Delta_x}, \frac{1}{\Delta_y}\right\} \quad (11)$$

When the spectrum is exactly not aliased, it is the optimal parameter we need

$$\beta_m = \min\left\{\frac{\pi\mathrm{m}}{2\Delta_x\left[W/2 - z_2\tan(\theta_y)\right]}, \frac{\pi\mathrm{m}}{2\Delta_y\left[W/2 - z_2\tan(\theta_y)\right]}\right\} \quad (12)$$

Given the same sampling intervals and viewing angles in the x and y directions, the following optimal parameter is designed

$$\begin{cases} R_m = \left[W/2 - z_2\tan(\theta_y)\right]/m, \\ \beta_m = \dfrac{\pi\mathrm{m}}{2\Delta_y\left[W/2 - z_2\tan(\theta_y)\right]} \end{cases} \quad (13)$$

OPM for aliasing-free reconstruction

Combining the frequency offset characteristics of FZA with the support domain constraint method, we can acquire aliasing-free spectral information for all apertures within the cutoff frequency. As a promising solution for optical superresolution, FP algorithm (*62*–*64*) has proven its effectiveness in FPM (*65*–*67*), long-range synthetic aperture (*68*, *69*), etc. In addition, we use the concept of a difference map (*70*–*73*) in ptychography and process multiple algorithmic steps in parallel to increase processing speed. Drawing on insights from FP about frequency-domain synthesis and difference map about parallel reconstruction, we can effectively restore all frequency information within $2/a \times 2/b$, achieving aliasing-free, high-resolution lensless reconstruction. The detailed OPM method is given in note S3 and fig. S2.

Imaging resolution analysis

For a lensless imaging system based on a programmable mask, the cutoff frequency is given by $f_c = m \times f_{\mathrm{LCD}}$, where $m = (z_1 + z_2)/z_1$ is the magnification and depends on the object-mask distance z_1 and the mask-sensor distance z_2. f_{LCD} is the cutoff frequency determined by the pixel size of the programmable mask. Ignoring the sensor's sampling limitations, the maximum spatial frequency that the system can resolve in the object space is $f_o = f_c \times z_2/z_1$. When using traditional reconstruction methods, the frequencies within f_o can be addressed, but ensuring imaging quality and SNR remains challenging.

The pixel pitch of the LCD screen used is 24 μm, and the pixel size is 9.6 μm, corresponding to a cutoff frequency of $f_{\mathrm{LCD}} = 1/f_{\mathrm{LCD}} = 1/(2\times 9.6\ \mu\mathrm{m}) = 52.08\ \ \mathrm{mm}^{-1}$. The USAF target used for the test measures 290 mm × 290 mm and is placed at a distance of $z_1 =$ 200 mm from the system. The distance between the LCD and the sensor is $z_2 = 1.66$ mm, with an equivalent magnification $m = (z_1 + z_2)/z_1 = 1.0083$. Therefore, the highest frequency response that can be received in the sensor plane is $f_c = m \times f_{\mathrm{LCD}} = 52.51\ \ \mathrm{mm}^{-1}$. The highest frequency response of the LCD, f_c, is lower than the highest frequency response of the sensor, $f_s = 1/[2\times(3\times 2.4\ \mu\mathrm{m})] = 69.44\,\mathrm{mm}^{-1}$, so the frequency information can be fully captured. The highest spatial frequency of the USAF target that can be captured is $f_o = f_c \times z_2/z_1 = 0.4358\ \ \mathrm{mm}^{-1}$, meaning the resolvable line-pair spacing (theoretical resolution) is $\Delta t = 2.29$ mm. The magnification of the tested USAF target compared to the standard USAF target is 3.8 times, corresponding to a line width of 314.98 μm × 3.8 = 1.20 mm for line pairs group 0 element 5, and a line-pair spacing (actual resolution) of $\Delta r = 1.20\ \mathrm{mm} \times 2 = 2.40$ mm. The relative error is $E = |\Delta t - \Delta r|/\Delta t = 4.8\,\%$, which mainly comes from the distance measurement error and the assembly error of the system.

Comparison of LIP and state-of-the-art lensless imaging methods

To illustrate the advantages of the FZA combined with our proposed algorithm over state-of-the-art (SOTA) lensless imaging methods, we chose separable mask (*13*), contour (*15*), diffuser (*14*), and random binary mask (*1*) as SOTA static-modulation masks for comparison, while selecting translated separable mask (*36*) and random binary array (*44*) for multishot method comparison.

For single-shot lensless methods, fig. S9A shows the PSF patterns corresponding to different static-modulation masks, where the FZA without offset and FZA with offset require the four-step phase shifting, and the FZA with OPM synthesizes the frequency response of various FZA offsets using our proposed OPM method. Figure S9B displays the normalized frequency response curves for each mask derived from the Fourier transform of each PSF pattern. The frequency responses of traditional static masks concentrate at the zero frequency, with a large direct current (dc) component, which suppresses responses at other higher frequencies. In contrast, the FZA-based masks, using the four-step phase-shift method, greatly reduce the impact of the dc component, allowing the high-frequency components to achieve better response. Notably, with our proposed OPM method, we can synthesize the parts of each subaperture that exhibit higher frequency responses, resulting in an almost flat response across the entire frequency range. More comparisons of LIP and static modulated lensless reconstruction results (*74*) are given in note S9 and fig. S10.

For multishot lensless methods, fig. S11 compares the FZA-based method with multishot methods, including translated separable mask from SweepCam (*36*) and random binary array from noise reduced dynamic synthetic coded aperture imaging camera (NoRDS-CAIC) (*44*). Figure S11 (A to D) shows each mask's patterns and the

corresponding normalized frequency responses. SweepCam uses translated separable masks to capture multiview information, while NoRDS-CAIC uses randomly distributed coded masks to obtain slightly varied responses. By comparison, although the frequency response of a single FZA mask is similar to the above two methods, with a higher response at zero frequency, the four-step phase-shift method effectively suppresses the influence of the dc component and improves response consistency. Likewise, when using our proposed OPM method, a uniform response is achieved across the entire frequency range, which is substantial for reconstruction quality. Figure S11E presents a comparison of reconstruction results between the multishot methods and the FZA-based method. Because of the limitations in frequency response, reconstruction results obtained with the translated separable mask and random binary array methods struggle to achieve accurate reconstruction under nonideal conditions, with limited improvement from multishot measurements. In contrast, the FZA-based method inherently provides strong noise robustness and further enhances resolution and SNR when combined with the OPM method. The qualitative comparison between LIP and existing static- and dynamic-modulation methods is given in note S9 and fig. S12.

In summary, the FZA-based approach offers an analytical mask form and is closely related to incoherent self-interference holography (*48*), which is grounded in established theoretical foundations. In contrast, traditional lensless imaging methods fundamentally model the measurements as the scene modulated by one or more masks, necessitating deconvolution for reconstruction. With only single-shot or random configurations, it is difficult to ensure stable deconvolution and reliable solutions from single or multiple measurements. Consequently, the FZA-based lensless imaging method enables a deeper understanding of the essence of lensless imaging, shifting the focus from an optimization problem to an optical problem with historical underpinnings. The unique features of the FZA-based lensless imaging, combined with the OPM method, effectively address the limitations encountered in SOTA lensless imaging methods, presenting a promising solution for high-quality, efficient lensless imaging. Moreover, by linking the underlying principles of incoherent self-interference holography (note S10 and fig. S13), we can rethink resolution analysis and aperture synthesis from a more optical perspective, enabling higher imaging quality and supporting more diverse imaging modes in lensless imaging.

Supplementary Materials

The PDF file includes:
Notes S1 to S10
Figs. S1 to S13
Legends for movies S1 to S4

Other Supplementary Material for this manuscript includes the following:
Movies S1 to S4

REFERENCES AND NOTES

1. R. Dicke, Scatter-hole cameras for X-rays and gamma rays. *Astrophys. J.* **153**, L101 (1968).
2. E. E. Fenimore, T. M. Cannon, Coded aperture imaging with uniformly redundant arrays. *Appl. Optics* **17**, 337–347 (1978).
3. E. Caroli, J. Stephen, G. Di Cocco, L. Natalucci, A. Spizzichino, Coded aperture imaging in X-and gamma-ray astronomy. *Space Sci. Rev.* **45**, 349–403 (1987).
4. A. Olivo, R. Speller, A coded-aperture technique allowing x-ray phase contrast imaging with conventional sources. *Appl. Phys. Lett.* **91**, 074106 (2007).
5. A. Greenbaum, W. Luo, T.-W. Su, Z. Göröcs, L. Xue, S. O. Isikman, A. F. Coskun, O. Mudanyali, A. Ozcan, Imaging without lenses: Achievements and remaining challenges of wide-field on-chip microscopy. *Nat. Methods* **9**, 889–895 (2012).
6. J. K. Adams, V. Boominathan, B. W. Avants, D. G. Vercosa, F. Ye, R. G. Baraniuk, J. T. Robinson, A. Veeraraghavan, Single-frame 3D fluorescence microscopy with ultraminiature lensless FlatScope. *Sci. Adv.* **3**, e1701548 (2017).
7. C. Zheng, D. Jin, Y. He, H. Lin, J. Hu, Z. Yaqoob, P. T. So, R. Zhou, High spatial and temporal resolution synthetic aperture phase microscopy. *Adv. Photonics* **2**, 065002 (2020).
8. X. Wu, J. Sun, J. Zhang, L. Lu, R. Chen, Q. Chen, C. Zuo, Wavelength-scanning lensfree on-chip microscopy for wide-field pixel-super-resolved quantitative phase imaging. *Opt. Lett.* **46**, 2023–2026 (2021).
9. Y. Xue, I. G. Davison, D. A. Boas, L. Tian, Single-shot 3D wide-field fluorescence imaging with a computational miniature mesoscope. *Sci. Adv.* **6**, eabb7508 (2020).
10. F. Tian, J. Hu, W. Yang, GEOMScope: Large field-of-view 3D lensless microscopy with low computational complexity. *Laser Photon. Rev.* **15**, 2100072 (2021).
11. X. Ye, X. Qian, Y. Chen, R. Yuan, X. Xiao, C. Chen, W. Hu, C. Huang, S. Zhu, T. Li, Chip-scale metalens microscope for wide-field and depth-of-field imaging. *Adv. Photonics.* **4**, 046006 (2022).
12. S. Utadiya, V. Trivedi, K. Bhanderi, M. Joglekar, C. Limberkar, K. Patel, G. Sheoran, H. Cabrera, B. Javidi, A. Anand, Thickness and surface profiling of optically transparent and reflecting samples using lens-less self-referencing digital holographic microscopy. *Appl. Surf. Sci. Adv.* **18**, 100484 (2023).
13. M. S. Asif, A. Ayremlou, A. Sankaranarayanan, A. Veeraraghavan, R. G. Baraniuk, Flatcam: Thin, lensless cameras using coded aperture and computation. *IEEE Trans. Comput. Imaging* **3**, 384–397 (2016).
14. N. Antipa, G. Kuo, R. Heckel, B. Mildenhall, E. Bostan, R. Ng, L. Waller, DiffuserCam: Lensless single-exposure 3D imaging. *Optica* **5**, 1–9 (2018).
15. V. Boominathan, J. K. Adams, J. T. Robinson, A. Veeraraghavan, Phlatcam: Designed phase-mask based thin lensless camera. *IEEE Trans. Pattern Anal. Mach. Intell.* **42**, 1618–1629 (2020).
16. N. Antipa, S. Necula, R. Ng, L. Waller. Single-shot diffuser-encoded light field imaging, in *IEEE International Conference on Computational Photography (ICCP)* (IEEE, 2016), pp. 1–11.
17. V. Boominathan, J. T. Robinson, L. Waller, A. Veeraraghavan, Recent advances in lensless imaging. *Optica* **9**, 1–16 (2022).
18. Y. Fan, J. Li, L. Lu, J. Sun, Y. Hu, J. Zhang, Z. Li, Q. Shen, B. Wang, R. Zhang, Q. Chen, C. Zuo, Smart computational light microscopes (SCLMs) of smart computational imaging laboratory (SCILab). *PhotoniX* **2**, 1–64 (2021).
19. J. Wu, H. Zhang, W. Zhang, G. Jin, L. Cao, G. Barbastathis, Single-shot lensless imaging with Fresnel zone aperture and incoherent illumination. *Light Sci. Appl.* **9**, 53 (2020).
20. Y. Ma, J. Wu, S. Chen, L. Cao, Explicit-restriction convolutional framework for lensless imaging. *Opt. Express* **30**, 15266–15278 (2022).
21. V. Boominathan, J. K. Adams, M. S. Asif, B. W. Avants, J. T. Robinson, R. G. Baraniuk, A. C. Sankaranarayanan, A. Veeraraghavan, Lensless imaging: A computational renaissance. *IEEE Signal Process. Mag.* **33**, 23–35 (2016).
22. A. Banerjee, H. Kumar, S. Saurav, S. Singh. Lensless image reconstruction with an untrained neural network, in *International Conference on Image and Vision Computing New Zealand* (Springer, 2022), pp. 430–441.
23. J. M. Bioucas-Dias, M. A. Figueiredo, A new TwIST: Two-step iterative shrinkage/thresholding algorithms for image restoration. *IEEE Trans. Image Process.* **16**, 2992–3004 (2007).
24. Z. Zhang, B. Zhang, X. Yuan, S. Zheng, X. Su, J. Suo, D. J. Brady, Q. Dai, From compressive sampling to compressive tasking: Retrieving semantics in compressed domain with low bandwidth. *PhotoniX* **3**, 19 (2022).
25. K. Monakhova, J. Yurtsever, G. Kuo, N. Antipa, K. Yanny, L. Waller, Learned reconstructions for practical mask-based lensless imaging. *Opt. Express* **27**, 28075–28090 (2019).
26. M. Lyu, H. Wang, G. Li, S. Zheng, G. Situ, Learning-based lensless imaging through optically thick scattering media. *Adv. Photonics.* **1**, 036002 (2019).
27. S. Feng, Q. Chen, G. Gu, T. Tao, L. Zhang, Y. Hu, W. Yin, C. Zuo, Fringe pattern analysis using deep learning. *Adv. Photonics.* **1**, 025001 (2019).
28. S. S. Khan, V. Sundar, V. Boominathan, A. Veeraraghavan, K. Mitra, Flatnet: Towards photorealistic scene reconstruction from lensless measurements. *IEEE Trans. Pattern Anal. Mach. Intell.* **44**, 1934–1948 (2020).
29. J. Wu, L. Cao, G. Barbastathis, DNN-FZA camera: A deep learning approach toward broadband FZA lensless imaging. *Opt. Lett.* **46**, 130–133 (2021).
30. A. Saba, C. Gigli, A. B. Ayoub, D. Psaltis, Physics-informed neural networks for diffraction tomography. *Adv. Photonics.* **4**, 066001 (2022).
31. F. Tian, W. Yang, Learned lensless 3D camera. *Opt. Express* **30**, 34479–34496 (2022).
32. J. Wu, V. Boominathan, A. Veeraraghavan, J. T. Robinson, Real-time, deep-learning aided lensless microscope. *Biomed. Opt. Express* **14**, 4037 (2023).
33. Q. Yang, R. Guo, G. Hu, Y. Xue, Y. Li, L. Tian, Wide-field, high-resolution reconstruction in computational multi-aperture miniscope using a Fourier neural network. *Optica* **11**, 860–871 (2024).
34. W. Chen, S. Feng, W. Yin, Y. Li, J. Qian, Q. Chen, C. Zuo, Deep-learning-enabled temporally super-resolved multiplexed fringe projection profilometry: High-speed kHz 3D imaging with low-speed camera. *PhotoniX* **5**, 25 (2024).

35. Y. Wu, M. K. Sharma, A. Veeraraghavan, WISH: Wavefront imaging sensor with high resolution. *Light Sci. Appl.* **8**, 44 (2019).
36. Y. Hua, S. Nakamura, M. S. Asif, A. C. Sankaranarayanan, Sweepcam–depth-aware lensless imaging using programmable masks. *IEEE Trans. Pattern Anal. Mach. Intell.* **42**, 1606–1617 (2020).
37. D. Deb, Z. Jiao, R. Sims, A. Chen, M. Broxton, M. B. Ahrens, K. Podgorski, S. C. Turaga, Fouriernets enable the design of highly non-local optical encoders for computational imaging. *Adv. Neural Inf. Process. Syst.* **35**, 25224–25236 (2022).
38. J. Qian, Y. Cao, Y. Bi, H. Wu, Y. Liu, Q. Chen, C. Zuo, Structured illumination microscopy based on principal component analysis. *eLight* **3**, 4 (2023).
39. J. R. Miller, C.-Y. Wang, C. D. Keating, Z. Liu, Particle-based reconfigurable scattering masks for lensless imaging. *ACS Nano* **14**, 13038–13046 (2020).
40. M. J. DeWeert, B. P. Farm, Lensless coded-aperture imaging with separable Doubly-Toeplitz masks. *Opt. Eng.* **54**, 023102 (2015).
41. T. Nakamura, T. Watanabe, S. Igarashi, X. Chen, K. Tajima, K. Yamaguchi, T. Shimano, M. Yamaguchi, Superresolved image reconstruction in FZA lensless camera by color-channel synthesis. *Opt. Express* **28**, 39137–39155 (2020).
42. Y. Wu, F. Li, F. Willomitzer, A. Veeraraghavan, O. Cossairt. WISHED: Wavefront imaging sensor with high resolution and depth ranging, in *IEEE International Conference on Computational Photography (ICCP)* (IEEE, 2020), pp. 1–10.
43. Y. Zheng, Y. Hua, A. C. Sankaranarayanan, M. S. Asif. A simple framework for 3D lensless imaging with programmable masks, in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)* (IEEE/CVF, 2021), pp. 2603–2612.
44. Z. Jiang, S. Yang, H. Huang, X. He, Y. Kong, A. Gao, C. Liu, K. Yan, S. Wang, Programmable liquid crystal display based noise reduced dynamic synthetic coded aperture imaging camera (NoRDS-CAIC). *Opt. Express* **28**, 5221–5238 (2020).
45. X. Chen, X. Pan, T. Nakamura, S. Takeyama, T. Shimano, K. Tajima, M. Yamaguchi, Wave-optics-based image synthesis for super resolution reconstruction of a FZA lensless camera. *Opt. Express* **31**, 12739–12755 (2023).
46. L. Mertz, N. Young, Fresnel transformations of images. *SPIE Milestone Series Ms* **128**, 44–49 (1996).
47. T. Shimano, Y. Nakamura, K. Tajima, M. Sao, T. Hoshizawa, Lensless light-field imaging with Fresnel zone aperture: Quasi-coherent coding. *Appl. Optics* **57**, 2841–2850 (2018).
48. J. Rosen, G. Brooker, Digital spatially incoherent Fresnel holography. *Opt. Lett.* **32**, 912–914 (2007).
49. K. Guo, S. Dong, G. Zheng, Fourier ptychography for brightfield, phase, darkfield, reflective, multi-slice, and fluorescence imaging. *IEEE J. Sel. Top. Quantum Electron.* **22**, 77–88 (2015).
50. C. Zuo, J. Sun, S. Feng, Y. Hu, Q. Chen, Programmable colored illumination microscopy (PCIM): A practical and flexible optical staining approach for microscopic contrast enhancement. *Opt. Lasers Eng.* **78**, 35–47 (2016).
51. J. Wu, F. Yang, L. Cao, Resolution enhancement of long-range imaging with sparse apertures. *Opt. Lasers Eng.* **155**, 107068 (2022).
52. N. Liu, Q. Zhao, G. Zheng, Sparsity-regularized coded ptychography for robust and efficient lensless microscopy on a chip, arXiv:2309.13611 [eess.IV] (2023).
53. L. Liu, W. Li, M. Gong, L. Zhong, H. Gu, S. Liu, Resolution-enhanced lensless ptychographic microscope based on maximum-likelihood high-dynamic-range image fusion. *IEEE Trans. Instrum. Meas.* **73**, 4502711 (2024).
54. Y. Zheng, S. Lin, C. Kambhamettu, J. Yu, S. B. Kang, Single-image vignetting correction. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**, 2243–2256 (2008).
55. S. J. Kim, M. Pollefeys, Robust radiometric calibration and vignetting correction. *IEEE Trans. Pattern Anal. Mach. Intell.* **30**, 562–576 (2008).
56. L. Li, S. Wang, F. Zhao, Y. Zhang, S. Wen, H. Chai, Y. Gao, W. Wang, L. Cao, Y. Yang, Single-shot deterministic complex amplitude imaging with a single-layer metalens. *Sci. Adv.* **10**, eadl0501 (2024).
57. Z. Zhou, Y. Zhang, Y. Xie, T. Huang, Z. Li, P. Chen, Y.-q. Lu, S. Yu, S. Zhang, G. Zheng, Electrically tunable planar liquid-crystal singlets for simultaneous spectrometry and imaging. *Light Sci. Appl.* **13**, 242 (2024).
58. C. Chen, H. Gu, S. Liu, Ultra-simplified diffraction-based computational spectrometer. *Light Sci. Appl.* **13**, 9 (2024).
59. X. Wu, J. Sun, Y. Chen, J. Wei, Q. Chen, T.-C. Poon, P. Gao, C. Zuo, Wavelength-scanning pixel-super-resolved lens-free on-chip quantitative phase microscopy with a color image sensor. *APL Photonics.* **9**, 016111 (2024).
60. K. Tajima, Y. Nakamura, K. Yamaguchi, T. Shimano, Improving resolution of lensless imaging with higher harmonics of Fresnel zone aperture. *Opt. Rev.* **29**, 153–158 (2022).
61. J. Sun, Q. Chen, Y. Zhang, C. Zuo, Sampling criteria for Fourier ptychographic microscopy in object space and frequency space. *Opt. Express* **24**, 15765–15781 (2016).
62. A. M. Maiden, M. J. Humphry, F. Zhang, J. M. Rodenburg, Superresolution imaging via ptychography. *J. Opt. Soc. Am. A* **28**, 604–612 (2011).
63. S. Dong, R. Horstmeyer, R. Shiradkar, K. Guo, X. Ou, Z. Bian, H. Xin, G. Zheng, Aperture-scanning Fourier ptychography for 3D refocusing and super-resolution macroscopic imaging. *Opt. Express* **22**, 13586–13599 (2014).
64. M. Stockmar, P. Cloetens, I. Zanette, B. Enders, M. Dierolf, F. Pfeiffer, P. Thibault, Near-field ptychography: Phase retrieval for inline holography using a structured illumination. *Sci. Rep.* **3**, 1927 (2013).
65. G. Zheng, R. Horstmeyer, C. Yang, Wide-field, high-resolution Fourier ptychographic microscopy. *Nat. Photonics.* **7**, 739–745 (2013).
66. J. Sun, Q. Chen, J. Zhang, Y. Fan, C. Zuo, Single-shot quantitative phase microscopy based on color-multiplexed Fourier ptychography. *Opt. Lett.* **43**, 3365–3368 (2018).
67. Y. Shu, J. Sun, J. Lyu, Y. Fan, N. Zhou, R. Ye, G. Zheng, Q. Chen, C. Zuo, Adaptive optical quantitative phase imaging based on annular illumination Fourier ptychographic microscopy. *PhotoniX* **3**, 24 (2022).
68. J. Holloway, M. S. Asif, M. K. Sharma, N. Matsuda, R. Horstmeyer, O. Cossairt, A. Veeraraghavan, Toward long-distance subdiffraction imaging using coherent camera arrays. *IEEE Trans. Comput. Imaging* **2**, 251–265 (2016).
69. J. Holloway, Y. Wu, M. K. Sharma, O. Cossairt, A. Veeraraghavan, SAVI: Synthetic apertures for long-range, subdiffraction-limited visible imaging using Fourier ptychography. *Sci. Adv.* **3**, e1602564 (2017).
70. V. Elser, Phase retrieval by iterated projections. *J. Opt. Soc. Am. A* **20**, 40–55 (2003).
71. P. Thibault, M. Dierolf, A. Menzel, O. Bunk, C. David, F. Pfeiffer, High-resolution scanning x-ray diffraction microscopy. *Science* **321**, 379–382 (2008).
72. P. Thibault, M. Dierolf, O. Bunk, A. Menzel, F. Pfeiffer, Probe retrieval in ptychographic coherent diffractive imaging. *Ultramicroscopy* **109**, 338–343 (2009).
73. L. Lu, J. Li, Y. Shu, J. Sun, J. Zhou, E. Y. Lam, Q. Chen, C. Zuo, Hybrid brightfield and darkfield transport of intensity approach for high-throughput quantitative phase microscopy. *Adv. Photonics.* **4**, 056002 (2022).
74. E. Agustsson, R. Timofte, NTIRE 2017 challenge on single image super-resolution: Dataset and study, *in Proceedings of the IEEE conference on Computer Vision and Pattern Recognition Workshops* (IEEE, 2017), pp. 126–135.

Acknowledgments

Funding: This work is supported by the National Natural Science Foundation of China [U21B2033 (C.Z.)], the National Major Scientific Instrument Development Project [62227818 (C.Z.)], the Leading Technology of Jiangsu Basic Research Plan [BK20192003 (Q.C.)], the Biomedical Competition Foundation of Jiangsu Province [BE2022847 (C.Z.)], the Key National Industrial Technology Cooperation Foundation of Jiangsu Province [BZ2022039 (C.Z.)], and the Fundamental Research Funds for the Central Universities [30920032101 (C.Z.)]. **Author contributions:** C.Z. and Q.C. initiated and supervised the project. C.Z., Q.C., X.Z., B.W., and H.G. conceived the idea and developed the methods. X.Z., B.W., and H.G. programmed the reconstruction algorithm and analyzed the data. X.Z., B.W., S.L., K.L., and H.G. conducted the simulations and experiments. All authors discussed the results and contributed to the writing of the paper. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 24 September 2024
Accepted 18 February 2025
Published 21 March 2025
10.1126/sciadv.adt3909