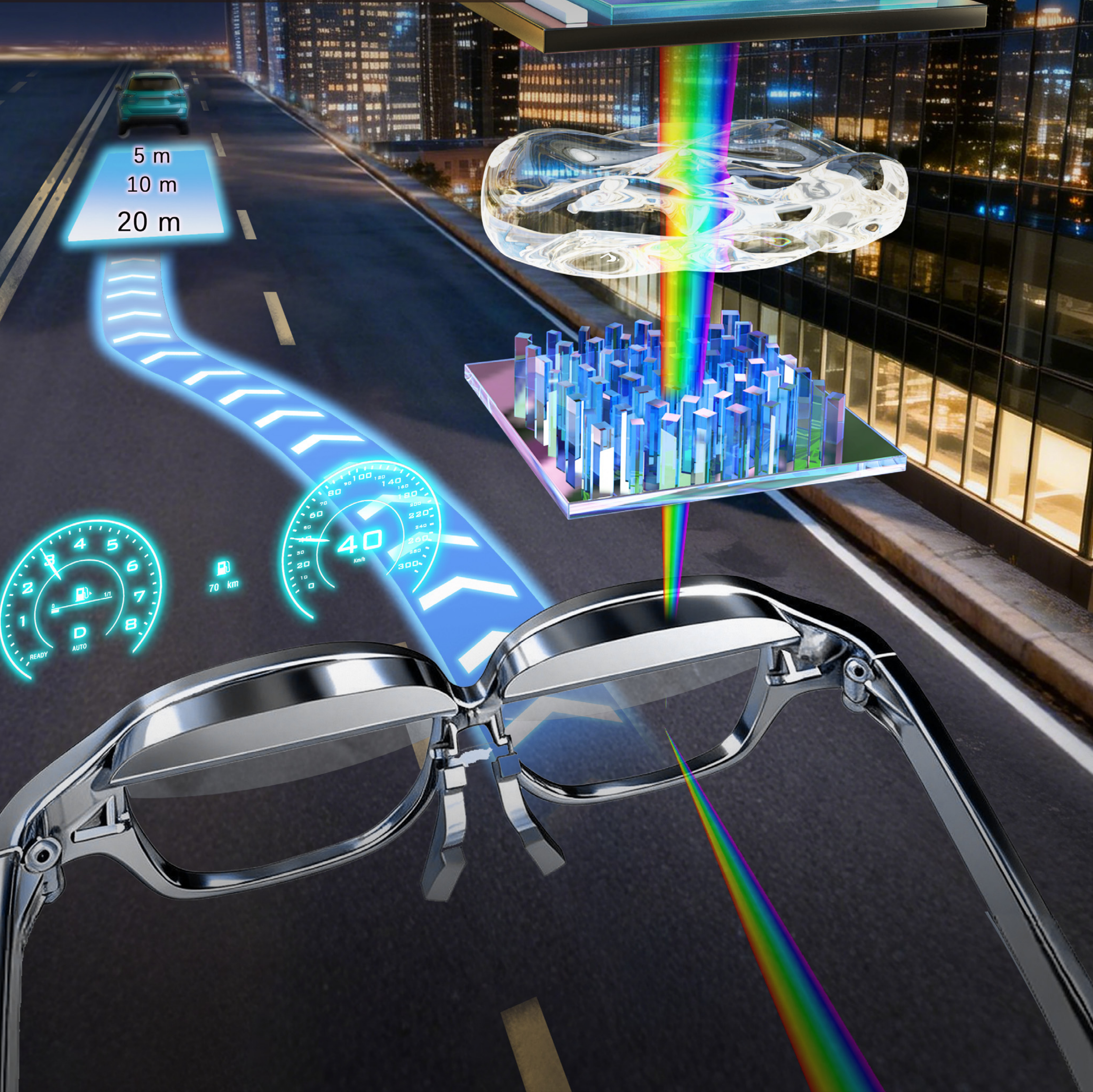
Opto-Electronic Science

Integrated metasurface-freeform system enabled multi-focal planes augmented reality display

Shifei Zhang†, Lina Gao†, Yidan Zhao, Yongdong Wang, Bo Wang, Junjie Li, Jiaxi Duan, Dewen Cheng, Cheng-Wei Qiu, Yongtian Wang*, Tong Yang* and Lingling Huang*

https://doi.org/10.29026/oes.2026.250031



Opto-Electronic Science

CN 51-1800/O4 ISSN (Print) 2097-0382 ISSN (Online) 2097-4000

Integrated metasurface-freeform system enabled multi-focal planes augmented reality display

Shifei Zhang, Lina Gao, Yidan Zhao, Yongdong Wang, Bo Wang, Junjie Li, Jiaxi Duan, Dewen Cheng, Cheng-Wei Qiu, Yongtian Wang*, Tong Yang* and Lingling Huang*

Citation: Zhang SF, Gao LN, Zhao YD, Wang YD, Wang B, Li JJ, Duan JX, Cheng DW, Qiu CW, Wang YT, Yang T, Huang LL. Integrated metasurface-freeform system enabled multi-focal planes augmented reality display. *Opto-Electron Sci* **5**, 250031(2026).

https://doi.org/10.29026/oes.2026.250031

Received: 1 August 2025; Accepted: 26 September 2025; Published online: 7 January 2026

Related articles

Ultracompact and high-efficiency liquid-crystal-on-silicon light engines for augmented reality glasses

Zhenyi Luo, Yuqiang Ding, Fenglin Peng et al

Opto-Electronic Advances 2024, **7**(10): 240039 doi: 10.29026/oea.2024.240039

IncepHoloRGB: multi-wavelength network model for full-color 3D computer-generated holography

Xuan Yu, Zhilin Teng, Xuhao Fan et al

Opto-Electronic Advances 2025, **8**(10): 250130 doi: 10.29026/oea.2025.250130

Eco-friendly quantum-dot light-emitting diode display technologies: prospects and challenges

Peili Gao, Chan Li, Hao Zhou et al

Opto-Electronic Science 2025, **4**(6): 240028 doi: 10.29026/oes.2025.240028

Metasurfaces for near-eye display applications

Yan Li, Xiaojin Huang, Shuxin Liu et al

Opto-Electronic Science 2023, **2**(8): 230025 doi: 10.29026/oes.2023.230025

More related articles in Opto-Electronic Journals Group website

Opto-Electronic Science

Article

January 2026, Vol. 5, No. 1

Open Access

Integrated metasurface-freeform system enabled multi-focal planes augmented reality display

Shifei Zhang[1,2†], Lina Gao[1,2†], Yidan Zhao[1,2], Yongdong Wang[1,2], Bo Wang[3], Junjie Li[3], Jiaxi Duan[1,2,4], Dewen Cheng[1,2], Cheng-Wei Qiu[5], Yongtian Wang[1,2]*, Tong Yang[1,2]* and Lingling Huang[1,2]*

Abstract: The advent of artificial intelligence (AI) has propelled augmented reality (AR) display technology to a pivotal juncture, positioning it as a contender for the next generation of mobile intelligent terminals. However, the pursuit of advanced AR displays, particularly those capable of delivering immersive 3D experiences, is significantly hindered by the performance limitations of current hardware and the complexity of system integration. In this study, we present an innovative multi-focal plane AR display system that integrates a non-orthogonal polarization-multiplexing metasurface, freeform optical elements, and an OLED display screen. All optical elements are integrated into a single solid-state architecture, based on a joint optimization design approach of ray tracing and diffraction theory. The multi-focal plane AR visual effect is realized by the compact and multiplexing metasurface, which performs distinct phase functions across diverse polarization channels. Meanwhile, freeform surfaces offer ample design flexibility for the collaborative optimization of multi-focal plane imaging and the see-through systems. Followed by a mechanical design and prototype assembly, we demonstrate the system's capabilities in real-time and multi-focal plane display. The digital images at all virtual image distances seamlessly integrate with the real environment, fully exhibiting the system's high parallelism and real-time interactivity. With the innovative design concept and joint design method, we believe that our work will spur more innovative and compact intelligent solutions for AR displays and inject new vitality into hybrid optical systems.

Keywords: augmented reality; metasurface-freeform; multi-focal planes display; non-orthogonal polarization-multiplexing; metasurfaces

DOI: 10.29026/oes.2026.250031 | CSTR: 32246.14.oes.2026.250031

Citation: Zhang SF, Gao LN, Zhao YD et al. Integrated metasurface-freeform system enabled multi-focal planes augmented reality display. *Opto-Electron Sci* **5**, 250031 (2026).

1 Introduction

Driven by the launch of artificial intelligence (AI), augmented reality (AR) display technology reaches a pivotal moment, and is widely expected to become the next generation of mobile intelligent terminals[1–5]. Numerous research efforts are dedicated to addressing the pressing issues of hardware limitations, mass production challenges, and enhancing the visual experience. However, the current approaches to real-time AR display technology are predominantly focused on single-focus plane displays, which lead to accommodation-vergence conflicts and are not conducive to long-duration use[6–8]. And existing solutions struggle to simultaneously meet the optical requirements for multi-depth projection, and the demands for compact form factor and affordable expense[2,9]. For instance, geometric optical schemes[10–12], which rely on the integration of multiple single-focus plane optical modules, encounter limitations including increased bulkiness and weight, and present

Received: 1 August 2025

Accepted: 26 September 2025

Published online: 7 January 2026

[1]School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China; [2]Beijing Engineering Research Center of Mixed Reality and Advanced Display, Beijing Institute of Technology, Beijing 100081, China; [3]Beijing National Laboratory for Condensed Matter Physics, Institute of Physics, Chinese Academy of Sciences, Beijing 100190, China; [4]Beijing NED+ AR Ltd., Beijing 100081, China; [5]Department of Electrical & Computer Engineering, National University of Singapore, Kent Ridge 117583, Singapore.

†These authors contributed equally to this work.

*Correspondence: YT Wang, E-mail: wyt@bit.edu.cn; T Yang, E-mail: yangtong@bit.edu.cn; LL Huang, E-mail: huanglingling@bit.edu.cn

certain challenges in terms of design and integration complexity. Additionally, schemes based on spatial light modulators (SLMs) and computer-generated holography (CGH) offer unique approaches for per-pixel depth control[1,13,14]. While, these technologies introduce high expenses and need large volume due to the requirements for laser sources and high resolution SLMs, and concerns regarding low efficiency and eye safety. The integral imaging schemes[15], which leverage microlens arrays to generate vivid 3D digital objects, face a trade-off between resolution and the number of viewpoints, as well as between field-of-view (FOV) and display quality, putting high requirements for micro-lens design and high-resolution micro-displays. To achieve the seamless integration of 3D or multi-focal plane digital scenes with real-world environments, novel advancements in optical material research, optical design methodologies, and process routes are imperative.

Recently, metasurfaces infuse optical display schemes with new vitality by offering sophisticated wavefront manipulations at the subwavelength scale[16,17]. Innovations such as achromatic metalens[18−22], meta-grating[23−26], metasurface holographic elements[27−30], as well as metasurface-based aberration compensator[31,32] microdisplays[33] and 3D sensors[34−36], are presented, driving the development of miniaturized and intelligent solutions to near-eye displays. However, the reported 3D or multi-focal plane display schemes typically utilize metasurfaces for encoding limited static 3D holograms[37−39], or employ them as meta-gratings for transmitting light modulated by preceding SLMs and lasers[1,13,14]. The SLMs, which functions as phase modulators, are not the mainstay in practical product applications; instead, devices that modulate light amplitude, such as micro display screens and liquid crystal on silicon (LCoS) devices are more widely utilized. Nevertheless, the implementation of a metasurface integrated with amplitude modulators for 3D AR display has rarely been explored. Furthermore, the wavefront multiplexing capability of metasurfaces remains largely untapped in creating multiple focal planes for real-time near-eye display. Despite the advancements in metasurface-based optical elements and modules, significant gaps remain in the systematic integration of multifunctional metasurface. These includes the joint design optimization methods of subwavelength diffractive and geometric optical elements, high precision and large-scale micro-nano manufacturing process, and prototype assembly techniques. And complete hybrid optical configuration, which simultaneously harnesses the strengths of metasurfaces in functionality multiplexing and geometric optical elements in optical path recirculation and relay functions, has not been presented and demonstrated.

To address these gaps, we conduct research on the design and demonstration of comprehensive hybrid optical system for real-time multi-focal plane AR display, which integrates an organic light-emitting diode (OLED) display screen, polarization-multiplexing metasurface, and freeform optical elements, as illustrated in Fig. 1(a). The system features a catadioptric birdbath structure, which is detailed in Fig. 1(b) and 1(c). The sufficient design freedom of metasurfaces and nonsymmetric freeform surfaces and the deep cooperation of these two kinds of elements ensures the system's increased imaging performance, compact size, lightweight nature and advanced AR display functions. Notably, the metasurface in our system serves as a polarization-multiplexing phase modulator, operating across various incident and fixed output polarization channels without altering the other optical elements. Such design maximizes the compactness, parallelism, and solid-state nature of the hybrid optical system. The display information with diverse polarization states is generated by different regions of the OLED display screen, which are affixed with various polarization films. After being processed by the metasurface, the information is manipulated to display at various depths, providing a 3D visual experience for customers (Fig. 1(d)). To significantly streamline the design process and conserve computational resources, we initially model the metasurface of diverse polarization channels as distinct phase functions within the ray tracing optimization framework. The optimization of multi-focal plane and see-through path configuration is conducted synergistically. The modulation transfer function (MTF) values of digital scenes and the see-through path exceed 0.25 at 20 lps/mm and 0.3 at 30 lps/mm, respectively, and the corresponding distortions are all well controlled. Subsequently, the non-orthogonal polarization-multiplexing phase functions are encoded by metasurface. The process differs from reported multiplexing of robust holography[40,41], and presents greater advances in reducing crosstalk and ensuring precision for specified wavefronts. Therefore, an established structural library with maximized design freedom and a machine learning framework are employed to maintain the accuracy of the phase functions as closely as possible. Followed by a mechanical design and prototype assembly, we demonstrate the system's capabilities in multi-focal plane display. All specified focal planes exhibit clear AR display effects, highlighting the real-time interactivity and feasibility of our approach. By fusing multi-focal plane digital information with the unrestricted physical environment, we believe our work would bring new possibilities in applications such as city navigation, entertainment, education, and industrial production.

2 Principles and methods

2.1 Multi-focal plane AR system design

A catadioptric birdbath (BB) system structure is used for the

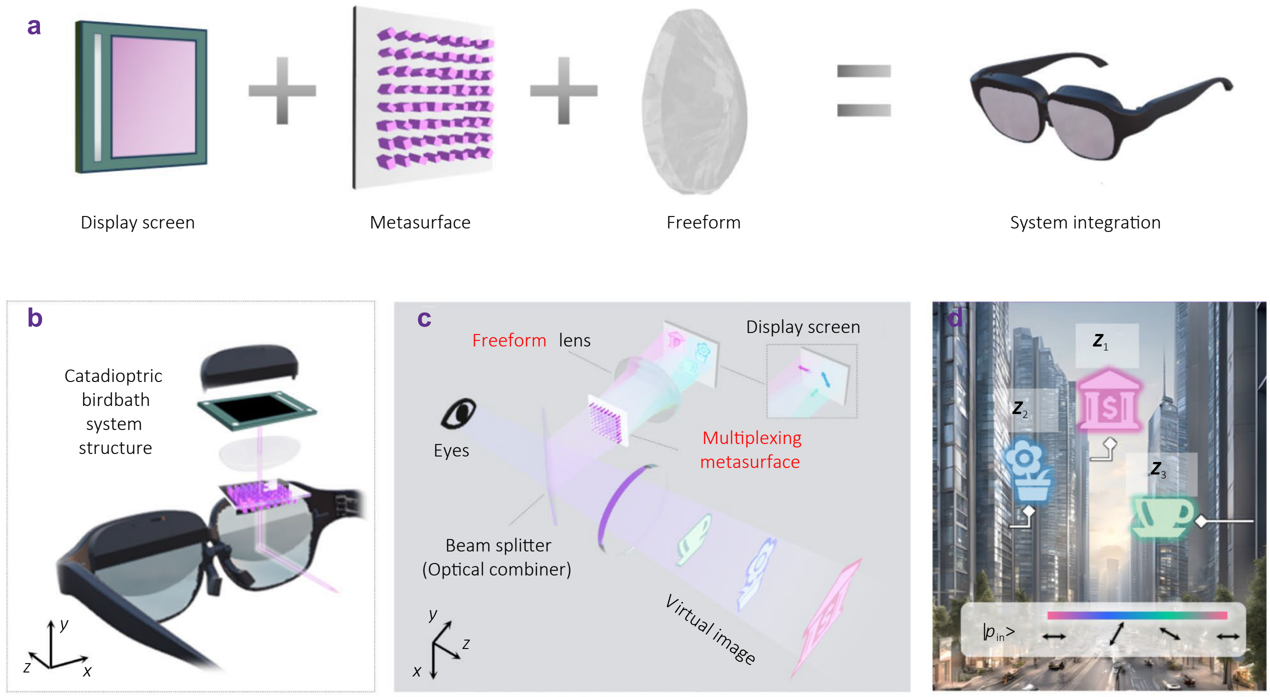


Fig. 1 | Schematic of the hybrid optical meta-freeform system for real-time 3D AR display. (a) The novel AR system integrates a display screen, a multiplexing metasurface, and a freeform lens simultaneously. (b) Catadioptric birdbath structure of the designed system. (c) Principle of polarization-multiplexing metasurface for multi-focal plane generation. The display screen is segmented into distinct regions, each equipped with a specific polarization film. (d) Applications of the 3D AR display in urban navigation.

design, as shown in Fig. 2. The light from the display screen travels through the polarization thin film, filter, lenses, metasurface and polarization analyzer, and is then reflected by the beam splitter (BS) and optical combiner (OC). Afterwards, the virtual information travels through the BS into the eye pupil, while the light from the outside scene travels through the OC and BS into the eye pupil directly. The light from different areas of the display screen forms different virtual image distances (VIDs) after travelling through the system, thus realizing multi-focal plane AR display. The VIDs of the system are 0.7 m, 1.5 m and 3 m. Because three different VIDs exist in the system, a multi-configuration system should be established during optical design with each configuration corresponds to a specific VID. The VID and the wavefront modulation effect of the metasurface vary with different configurations, while the shape and locations of other optical elements are the same. Considering the challenges of large-area, high-precision fabrication of the metasurface, reducing its size facilitates manufacturing. To reduce the size of metasurface, the light beams of different fields and configurations should overlay on the metasurface as much as possible, which means the metasurface should be located at the conjugate position of aperture stop inside the system. To realize this structure, an intermediate image will appear, which is different from traditional BB structure. For the optical design of visual or AR/VR system, a reverse optical path is generally used, where the object in the reverse system is the virtual image plane in the actual system, and the image plane in the reverse system is the display screen. The aperture stop (eye pupil in the actual system) is located at the beginning of the whole system. A 0.71-inch (1 inch = 2.54 cm) OLED display with resolution of 1920 × 1080 pixels is selected. The material of lenses is K22R. The material of the metasurface substrate is SiO_2 and the thickness is 0.5 mm.

Different polarization channels for different VIDs are multiplexed in a single system, with each configuration using a specific non-central region of the screen, Therefore, the optical path of each configuration is off-axis nonsymmetric, and thus induce severe non-symmetric and unconventional field-dependence aberrations, which are difficult to be corrected using traditional rotationally-symmetric spheres and aspheres. Off-axis asphere has nonrotationally symmetric surface shape and can be considered a kind of freeform surface. However, as the off-axis asphere is part of its parent rotationally symmetric aspherical surface, the degrees of design freedom of an off-axis asphere is only one more (the off-axis distance) compared with its parent asphere[42]. In other words, the coefficients of the nonsymmetric terms in an off-axis asphere are interrelated. To breakthrough these limitations, in this work, true freeform optical element (the coefficients of the nonsymmetric terms are independent compared with off-axis aspheres) is

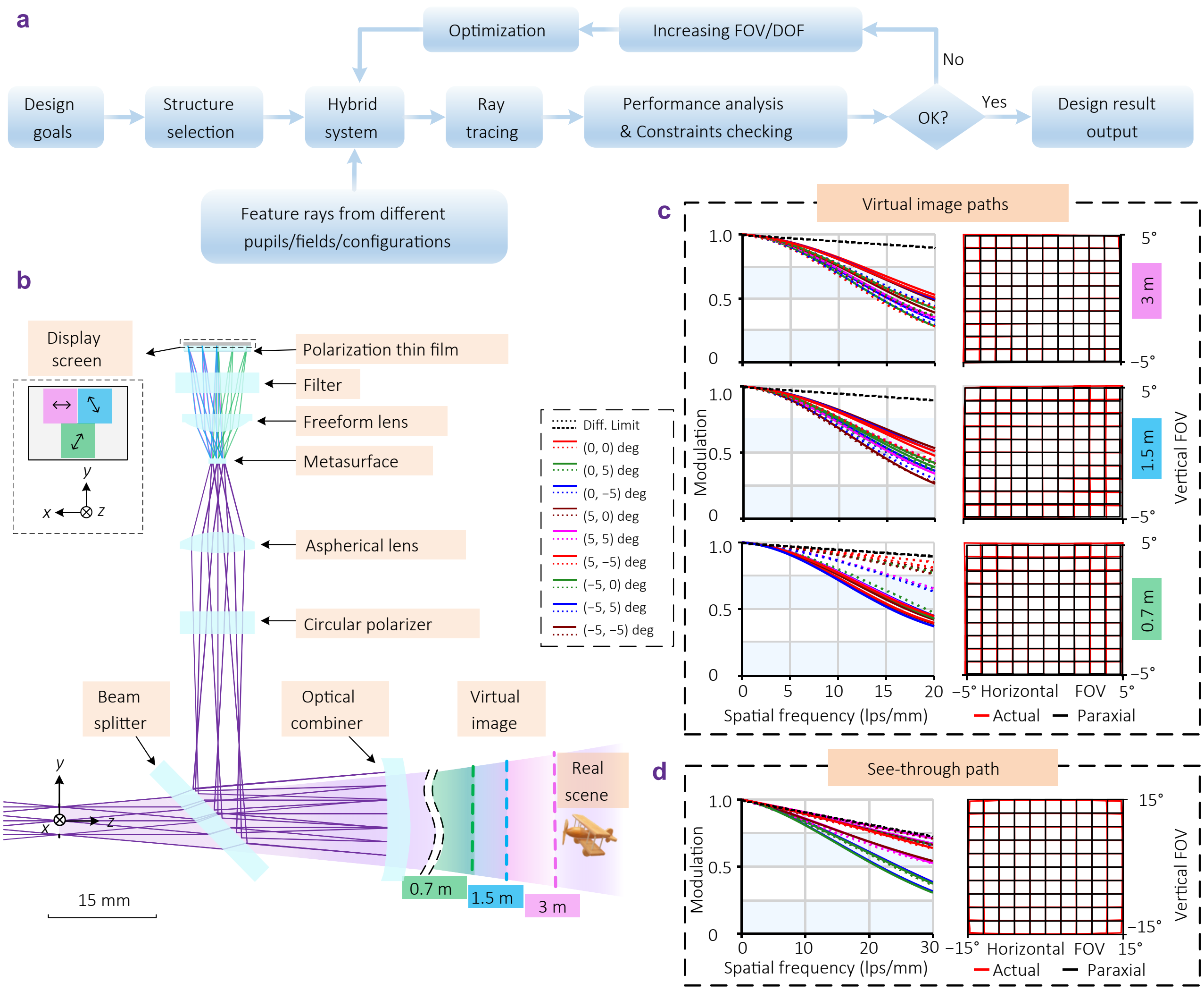


Fig. 2 | Multi-focal plane AR display system design. (a) Optical design framework for the system. (b) The layout of the system. (c) MTF plots and distortion grids for the virtual image light paths. (d) MTF plot and distortion grid for the see-through light path.

induced into the system. In comparison with traditional spherical or aspherical surfaces, freeform optical surface has no rotationally symmetry[43−45], and has much more parameters in surface description. The increase of design freedom and structure flexibility can dramatically improve the imaging performance of the system, while reduce the system size, weight, and number of elements. In addition, the aberrations induced by freeform surfaces can well compensate the aberrations induced by non-symmetric system structure in theory[46]. As a result, combining polarization-multiplexing metasurface and freeform optics can lead to the successful design of multi-focal plane AR display.

For the design of imaging optics, multi-parameter optimization is generally used to get a good imaging performance. Freeform surface is a kind of "pure geometric optical element", which can be modeled and handled by traditional optical design theory. However, for metasurface, it is a multifunctional diffractive optical element with subwavelength sampling spacing. For the design of imaging system with electromagnetics theory, the optimization and analysis process will be very complicated and very time- and resource-consuming. To solve this problem, here we consider that the wavefront modulation effect of metasurface can be characterized by its phase function $\phi(x, y)$. The ray propagation at metasurface can be modeled by generalized refraction or reflection law, where the gradient of $\phi(x, y)$ can be easily calculated. In this way, the design and optimization of imaging systems containing both geometric optical elements (such as freeform optics) and metasurface can be unified into the ray tracing approach.

A successive optimization strategy is employed to design the system. First, only the central 0° field is used for each configuration. At this stage, the surface type of the lenses and OC are all spheres. The chief ray of each central field is optimized to be intersect with the image plane at the center of each display area. As the light of three configurations should be deflected to specific areas on the image plane, the phase function $\phi(x, y)$ of three configurations should be all nonsymmetric, or can be seen as freeform. For the mathematical expressions of the nonsymmetric phase functions and the nonsymmetric freeform surface, there are many choices such as *XY* polynomials, Zernike polynomials, Forbes *Q*-polynomials, and NURBS. In this paper, we choose *XY* polynomials as the mathematical expressions. Compared to other expressions, *XY* polynomials have a simpler description, and is expressed in commonly used Cartesian coordinates. This kind of expressions is generally directly included in common fabrication machines, and therefore is more compatible with fabricating. In addition, for Zernike polynomials and Forbes *Q*-polynomials, their special orthogonal properties exist in a unit circular domain, which is not consistent with the proposed design as square apertures are used for the freeform element incorporating an overall rectangular light footprint of the entire rectangular field-of-view. The NURBS surface or phase function representations require substantial iterative process for the ray tracing process and will significantly affect the design efficiency. Therefore, in this design, we use *XY* polynomials as the mathematical expressions for the metasurface and the freeform surface. In addition, one of the design principles for freeform optical design is that the designers should use as few and low orders freeform terms as possible to complete the system design. Using unnecessary higher orders terms will decrease the design efficiency, increase the fabrication difficulty, and may make the optimization unstable. Based on the aberration theory, polynomials up to the 8th order corresponds to the dominant 8th order wave aberrations to a great extent[47]. Using higher order terms has negligible impact on aberration correction in this system. Therefore, in this design, the *XY* polynomials up to 8th order ($m+n\leq 8$) are used for both the freeform surface and the metasurface. Here, the freeform phase function of the qth configuration $\phi_q(x, y)$ is described as

$$\phi_q(x,y)=\frac{2\pi}{\lambda_c}\sum_{m=0}^{8}\sum_{n=0}^{8}A_{m,n,q}x^m y^n \quad (1\leq m+n\leq 8), \quad (1)$$

where λ_c is the working wavelength, and $A_{m,n,q}$ is the coefficients of $x^m y^n$ phase terms of the qth configuration. (x, y) are the local coordinates of the substrate surface. Then, the FOV of each configuration is gradually increased. The feature rays from different pupil coordinates, different fields and different configurations should be sampled during optimization. As the ideal image point of each field can be calculated based on the object-image mapping relationship, distortion and the magnification of the system are controlled by constraining the real ray trace data of the marginal fields in each configuration. To correct the aberrations, more degrees of freedom (DOF) for design are introduced gradually. The surface types of the lenses are gradually upgraded to aspheres, and two surfaces are further upgraded to freeform surface. The freeform surface type is *XY* polynomials surface, whose surface expression can be written as:

$$z(x,y)=\frac{c\,(x^2+y^2)}{1+\sqrt{1-(1+k)c^2\,(x^2+y^2)}}+\sum_{m=0}^{8}\sum_{n=0}^{8}C_{m,n}x^m y^n \quad (1\leq m+n\leq 8), \quad (2)$$

where c is the curvature, k is the conic constant, $C_{m,n}$ is the coefficients of high-order terms, (x, y) is the local coordinates of the substrate surface, and z is the surface sag. The size of the metasurface is controlled by constraining the ray intersection with the metasurface of the marginal pupil rays. During optimization, the image performance of the see-through path should also be maintained and the distortion should be controlled. This is realized by adding a fourth configuration during the design and optimizing the surface shape of the outside surface of the OC simultaneously. The above is the system design process, as shown in Fig. 2(a).

After all optimization steps are completed, the final system is obtained, the structure is shown in Fig. 2(b). The size of the metasurface is 2 mm × 2 mm. For the virtual image light paths, the full FOV of the system is 10°×10° and the entrance pupil diameter is 4 mm. The working wavelength is 632.8 nm and the bandwidth is 10 nm. Figure 2(c) shows the modulation transfer function (MTF) curves and distortion grids. The MTF curves of 9 typical fields, including the center and marginal fields, are provided. In the optimization process, the weights of different field points are adjusted to make the image quality across the full FOV as balanced as possible. All the MTF values for the three configurations are higher than 0.25 at 20 lps/mm. The distortion of the system is well controlled. Good imaging performance is achieved.

For the see-through light path working at the visible spectral range, 30°×30° FOV is considered during the design. MTF plot and distortion grid is shown in Fig. 2(d). The MTF values for all fields are higher than 0.3 at 30 lps/mm. The distortion of see-through light path is well controlled. Microstructures on metasurfaces are designed based on phase functions.

2.2 Metasurface design

The optimized phase functions corresponding to the three multi-focal planes are encoded within the distinct polarization channels of the anisotropic metasurface, denoted as ϕ_1, ϕ_2, ϕ_3. To achieve a robust, rapid, and compact AR display system, we eschew the use of mechanical or electrical polarization switching mechanisms, which are not applicable in realistic cases. Instead, the polarization-multiplexing metasurface employs a fixed analyzer at the exit end, thereby eliminating the need for dynamic polarization switching for the output. Concurrently, polarization thin films with distinct linear polarization states are affixed to specific regions of the display screens to generate diverse polarized incidences directed towards the metasurface. Here, the anisotropic metasurface provides an optical response characterized by a Jones matrix $\boldsymbol{J}$. The incident polarization states are defined as $|P_q\rangle = \left[\cos\left(\frac{q-1}{3}\pi\right), \sin\left(\frac{q-1}{3}\pi\right)\right]^{\mathrm{T}}$ for $q = 1, 2, 3$, and right circular polarization (RCP) is employed as the analyzer. This design ensures a near-equal distribution of energy across the multiplexing channels. Consequently, the encoded modulation fields for these polarization channels are represented as $\langle R|\boldsymbol{J}|P_q\rangle$ (Fig. 3(a)). And this design exhibits a purely solid-state and compact architecture, ensures a high-frame-rate, multi-focal plane AR display without reliance on time-multiplexing.

However, precision wavefront encoding based on non-orthogonal polarization channels is challenging to achieve, as crosstalk inevitably exists for non-orthogonal polarization channels, in accordance with Malus' law. To address this issue, we optimize the encoding process of phase functions through the metasurface design and encoding algorithm simultaneously. Firstly, as depicted in Fig. 3(b), we adopt a diatomic strategy for the metasurface design. The synthesized sampling units consist of two distinct types of meta-atoms arranged in a staggered distribution. These meta-atoms, characterized by varying lengths, widths, and rotation angles (denoted as l_1, w_1, θ_1, l_2, w_2, and θ_2), provide the optical response of distinct unitary matrices $\boldsymbol{J}_1$ and $\boldsymbol{J}_2$. The Jones matrices for each type of meta-atom are mathematically formulated as $\boldsymbol{J}_s = \exp(\mathrm{i}\varphi_{s1})|\theta_s\rangle\langle\theta_s| + \exp(\mathrm{i}\varphi_{s2})|\theta_{s,\perp}\rangle\langle\theta_{s,\perp}|$ for s=1,2. $|\theta_s\rangle$ and $|\theta_{s,\perp}\rangle$ represent the polarization eigenstates of $\boldsymbol{J}_S$, with $|\theta_s\rangle$=[cosθ_s, sinθ_s]$^{\mathrm{T}}$; exp(iφ_{s1}) and exp(iφ_{s2}) denote the eigen-phase terms associated with each polarization state. In this case, the Jones matrix of the synthesized unit can be represented as the average of $\boldsymbol{J}_1$ and $\boldsymbol{J}_2$: $\boldsymbol{J} = (\boldsymbol{J}_1+\boldsymbol{J}_2)/2$. With this approach, the full design freedom of $\boldsymbol{J}$ based on single-layer design can be achieved, given that the eigen-phase terms of $\boldsymbol{J}_1$ and $\boldsymbol{J}_2$ cover a complete phase

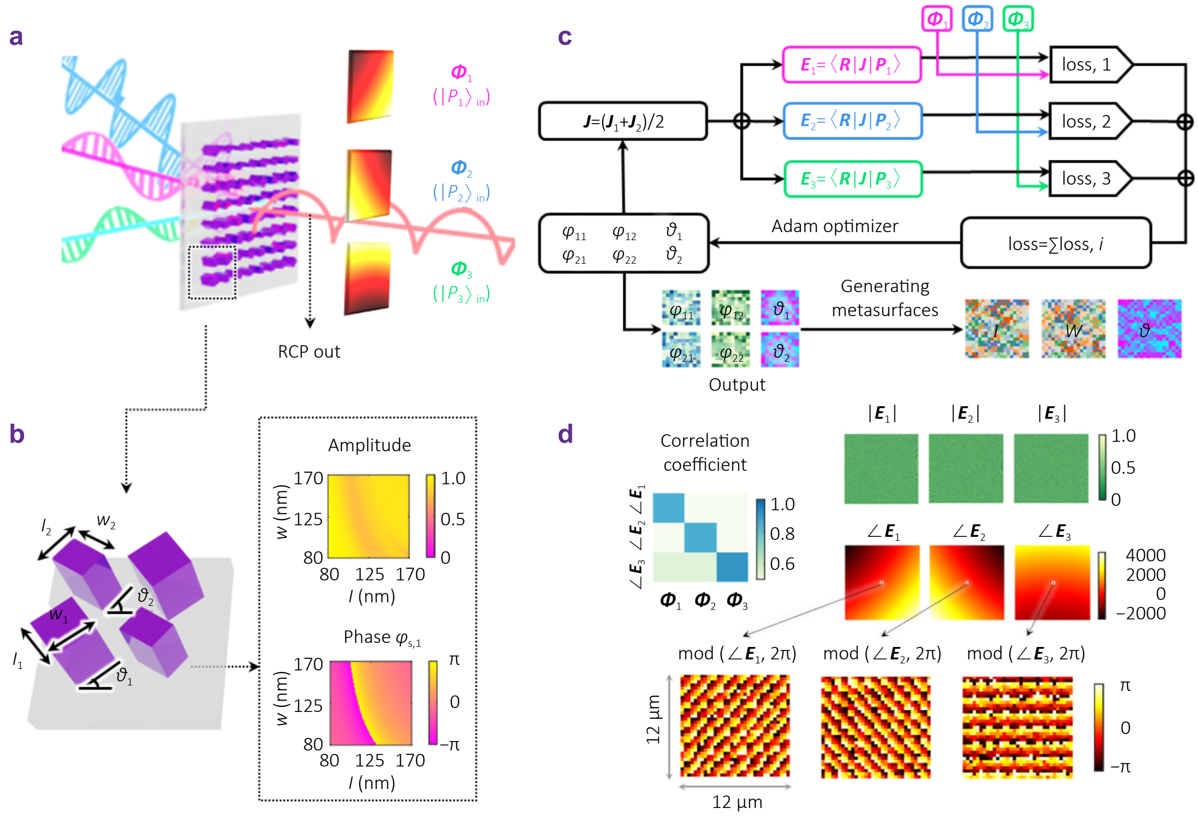


Fig. 3 | Metasurface design and optimization for multifunctional phase modulation. (a) Schematic of a polarization-multiplexing metasurface designed for phase encoding. (b) A diatomic design strategy for comprehensive Jones matrix modulation, featuring six adjustable structural parameters. (c) Optimization flowchart for the phase encoding based on the polarization-multiplexing metasurface. (d) Reconstructed results of polarization channels, including amplitude and phase distributions, and the correlation coefficient matrix for multiplexed channels.

cycle from 0 to 2π (further proof is provided in the supplementary information).

Here, we select the amorphous silicon nanofins as the meta-atoms, with their period and height configured as 300 nm and 400 nm, respectively. Moreover, the period of synthesized units remains smaller than the wavelength, facilitating precise wavefront at a working wavelength of 632.8 nm. Using the rigorous coupled-wave analysis (RCWA) method, we sweep through the optical responses of nanofins with diverse lengths and widths, ranging from 80 nm to 170 nm. Simulations reveal that the nanofins exhibit uniform amplitude and phase modulation covering 0 to 2π along both long and short axis directions. The optical responses along the long axis are presented in Fig. 3(b), while those along the short axis are transposed, given the interchangeability of length and width in nanofins. To conclude, the Jones matrix of this single-layer structure is endowed with maximum design freedom.

Additionally, the metasurface design is optimized within a machine learning framework that employs the Adam optimizer (Fig. 3(c)). The optimization initiates with the random generation of metasurfaces, driven by input parameters φ_{11}, φ_{12}, θ_1, φ_{21}, and θ_2. For each polarization channels, the modulated light field $\langle R|\boldsymbol{J}|P_q\rangle$ is compared with the respective target phase functions $\exp(\mathrm{i}\varphi_q)$. The loss functions involve the mean square errors (MSE) of both amplitude and phase distributions. Through iterative optimization, the parameters are refined to achieve low MSE values as low as 0.038 for amplitude and 0.504 rad for phase across all channels. The correlation coefficients, which quantify the resemblance between the reconstructed and target light fields, exceed 0.8 for each polarization channel. Furthermore, the magnified view in Fig. 3(d) reveals the wrapped phase corresponding to each polarization state, demonstrating the metasurface's ability to independently encode each channel with high fidelity. Based on the results, the layout of the metasurfaces is meticulously constructed through a pixel-by-pixel selection process, utilizing meta-atoms from the established structural library.

3 Prototype and experimental results

The metasurface, with a footprint of 4 mm^2, is fabricated employing a standard silicon-based semiconductor process. The sequence mainly includes initial film deposition via plasma enhanced chemical vapor deposition (PECVD), high-resolution pattern transfer using electron beam lithography (EBL), the lift-off process to transfer the pattern from the resist to a chromium mask, and finally, inductively coupled plasma reactive ion etching (ICP-RIE) to achieve

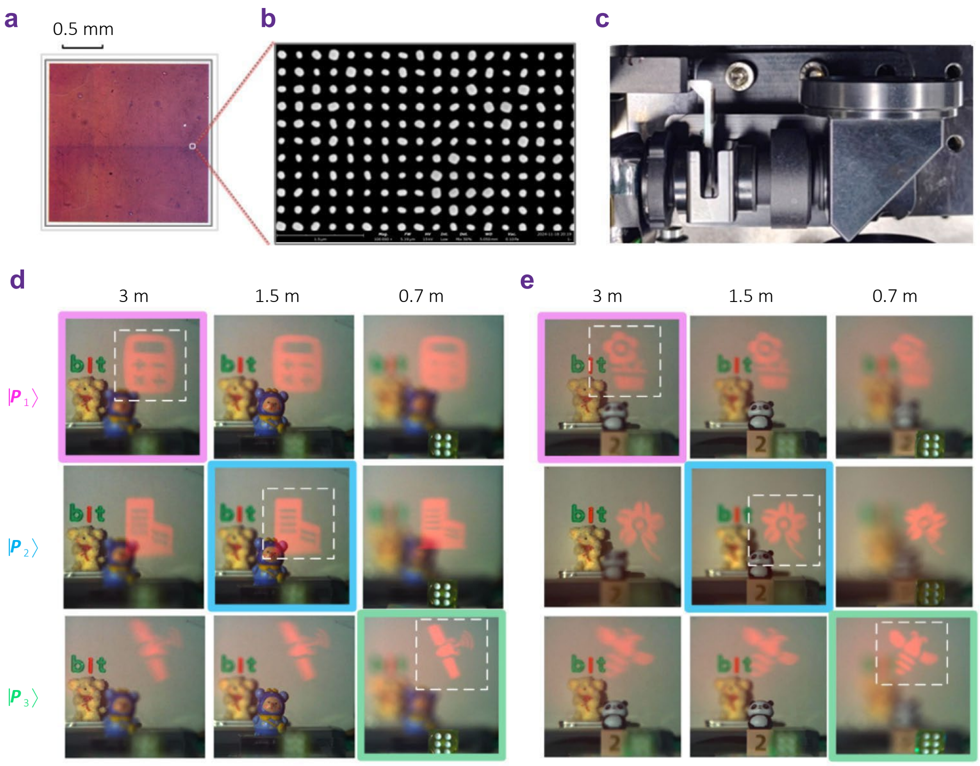


Fig. 4 | Prototype and experimental outcomes of the multi-focal plane AR display system. (a) Microscopic image and (b) scanning electron microscopy (SEM) image of the fabricated metasurface. (c) The AR display system prototype. (d, e) The experimental results of image capture for a single display area.

the metasurface's intricate structures. Figure 4(a, b) shows microscopic image and scanning electron microscope (SEM) images of the fabricated metasurfaces. The lenses are fabricated using single-point diamond turning (SPDT). Both the freeform and aspherical lenses are coated with an anti-reflection coating, and the OC is coated with a semi-transparent and semi-reflective coating. Additionally, a mechanical system is also designed for the assembly and alignment of geometric optical elements and metasurfaces (Fig. 4(c)). Considering the inevitable manufacturing and assembly errors, we also allocate an adjustment margin for the display screen and metasurface to further enhance the system's performance. The tolerance analysis results of the system and the system alignment process are provided in the supplementary information. For validation, a camera is placed at the exit pupil of the optical system to replace the human eye. The optical axis of the camera is aligned with the viewing axis of the system.

Firstly, we validate the AR display effect for each polarization channel at respective VID, both the real and digital scene can be seamlessly transitioned in real time. The experimental results of imaging are shown in Fig. 4(d). In the experimental setup, three real objects are placed at different distances from the camera within a real-world scene: a yellow puppy at 3 m, a blue bear at 1.5 m, and a green dice at 0.7 m. The display areas corresponding to the VIDs of 3 m, 1.5 m, and 0.7 m present the calculator, house, and satellite patterns, respectively. These patterns are displayed individually, with each appearing sharp (highlighted in colors that correspond to their respective polarization channels) at its designated VID and becoming progressively blurred when the focus shifts to other depths. For each virtual image, valid focusing and blurring effects are clearly observed. Figure 4(e) further validates the functionality of the system by showcasing another scenario with actual objects and digital objects.

Subsequently, we assess the multi-focal plane display capabilities of our AR system at three distinct distances simultaneously, which fully exhibit the system's high parallelism and real-time interactivity. To enhance observability, the final display icons are strategically offset planarly, as depicted in Fig. 5. Each VID's corresponding projected images are outlined by a white frame, presenting sharp and detailed information. As the focus shifts, the images gradually blur, and demonstrate the precision in depth perception, which is crucial for enhancing the user's spatial awareness and interaction with the environment. The seamless integration of diverse digital to overlay on the physical environment

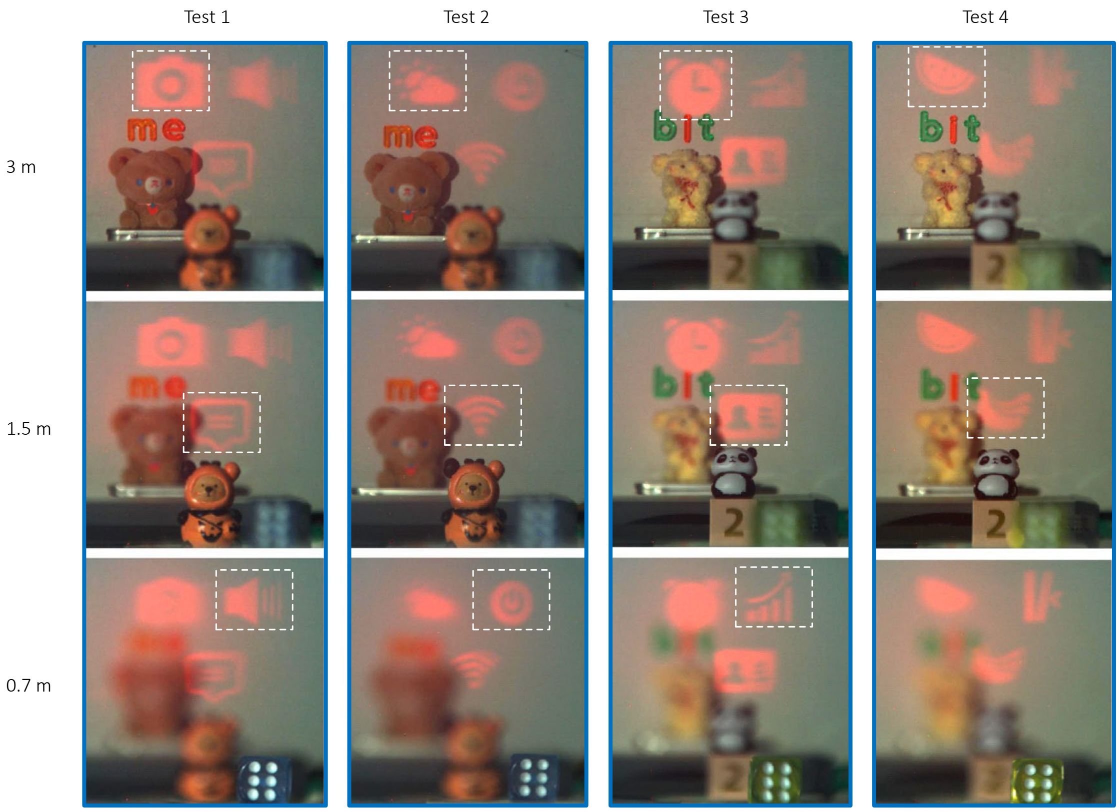


Fig. 5 | Multi-focal plane display of our prototype at three distinct distances simultaneously. The image that is outlined by a white frame correspond to different VID. Both the real and digital scene can be seamlessly transitioned in real time.

underscores the system's adaptability. The digital icons, representing various applications, at different spatial locations signifies the potential for spatial interaction that transcends traditional depth limitations. The interaction of fruit icons with real-world objects like bears and pandas suggests potential applications in AR gaming and education. To further validate the system's capabilities, we generate a dynamic three-dimensional AR display video, which is provided in the supplementary information.

4 Conclusions

The experiments successfully validate the feasibility of our hybrid meta-freeform optics architecture and joint design approach. And the approach strategically harnesses the distinct advantages of each optical element. Particularly, the polarization-multiplexing strategy efficiently integrates multiple functions onto a lightweight and compact metasurface, just millimeters in size, without altering other optical elements. This ingenious strategy greatly reduces the system's volume while maintaining special functionality, and eliminates the need for complex multi-system integrations. The introduction of freeform surfaces offers ample design flexibility for the collaborative optimization of multi-focal imaging and the see-through systems. To our knowledge, no prior prototype yet to be reported previously that integrated a display screen, metasurface, and freeform optical elements within a single comprehensive optomechanical design for multi-focal plane AR display.

Furthermore, to meet the demands of a purely solid-state optical tunable system, non-orthogonal polarization multiplexing channels are employed. Recently, strategies involving noise engineering[41] and algorithm optimization[40] have been proposed to achieve high-fidelity holographic encoding across various polarization channels. However, holography is inherently robust to noise, allowing holographic images to remain visually discernible even when there is a significant drop in the signal-to-noise ratio (SNR). To our knowledge, the demonstration of accurate encoding for phase functions has not been experimentally achieved.

Due to the significant size disparity between the display screen and the metasurface, additional optical elements and an intermediate image plane are incorporated, leading to the current prototype's dimensions of 9.3 cm×4.5 cm×4.9 cm. However, with the advent of smaller display screens, such as 0.13-inch Micro LED micro-display[48], and a corresponding slight increase in metasurface dimensions, a more compact optical system can be achieved.

Furthermore, more focus planes can be customized on this basis, facilitating a more continuous 3D AR display effect. It should be noted that, there is an inherent tradeoff between the number of display planes and critical optical performance factors such as the efficiency of each polarization channel, the accuracy of encoded wavefronts, and imaging resolution. Despite this, considering the advancements of high-resolution display products and metasurface design, the improvements in axial resolution are not only feasible but also unstoppable.

Full-color, which is also a crucial metric for near-eye display, is also promising by further optimization of metasurface design and effective integration with full-color display screen. By employing physical mechanisms such as bound states in the continuum (BIC)[26,49] and dispersion engineering[50], along with cutting-edge design methods including inverse design and AI-driven frameworks, the wavefront multiplexing of metasurfaces across diverse RGB and polarization channels is becoming increasingly feasible. Simultaneously, it is crucial to consider the characteristics of the display screen within the overall system design, including pixel arrangements and the division of distinct regions, efficiency and the bandwidth of RGB units.

In this paper, the design and optimization of the hybrid imaging system is unified into the ray tracing approach in order to significantly accelerating the design process and reduce the computational cost. Designing such system using electromagnetics theory considering the true interaction between the light and the metasurface as well as geometrical lenses is possible using existing theories, but may be very complicated and very time- and resource-consuming. In future works, besides using new and much faster design and simulation methods, differential design framework can also be explored. The wavefront modulation effect of the metasurface is no longer characterized using a phase function $\phi(x, y)$, but is directly characterized by the light matter interactions indued by specific meta-atoms. The interactions of the light (different directions, polarization, wavelength) with the meta-atoms (with different size, height, materials, etc.) are directly integrated into the design process as a specific mathematical function. In this way, the system process combining ray tracing may consider the light-matter interactions and electromagnetics theory, andcan enable more powerful inverse design of such hybrid systems.

To conclude, our work provides a compact, high-framerate, real-time, and pure-solid-state solution for multi-focal plane AR display applications. The innovative system design concept offers a distinctive alternative for state-of-the-art display and imaging technologies. The combined approach of ray tracing and diffraction theory conserves both time and computational resources, provides an invaluable reference for the practical application of metasurfaces and the design of future hybrid optical systems.

References

1. Gopakumar M, Lee GY, Choi S et al. Full-colour 3D holographic

augmented-reality displays with metasurface waveguides. *Nature* **629**, 791–797 (2024).
2. Xiong JH, Hsiang EL, He ZQ et al. Augmented reality and virtual reality displays: emerging technologies and future perspectives. *Light Sci Appl* **10**, 216 (2021).
3. Hua H, Javidi B. Augmented reality: easy on the eyes. *Opt Photonics News* **26**, 26–33 (2015).
4. Azuma RT. A survey of augmented reality. *Presence Teleoperators Virtual Environ* **6**, 355–385 (1997).
5. Kress BC. *Optical Architectures for Augmented-, Virtual-, and Mixed-Reality Headsets* (SPIE Press, Bellingham, 2020).
6. Hoffman DM, Girshick AR, Akeley K et al. Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. *J Vis* **8**, 33 (2008).
7. Kooi FL, Toet A. Visual comfort of binocular and 3D displays. *Displays* **25**, 99–108 (2004).
8. Ding YQ, Gu YC, Yang Q et al. Breaking the in-coupling efficiency limit in waveguide-based AR displays with polarization volume gratings. *Light Sci Appl* **13**, 185 (2024).
9. Park JH, Lee B. Holographic techniques for augmented reality and virtual reality near-eye displays. *Light Adv Manuf* **3**, 137–150 (2022).
10. Wilson A, Hua H. Design and demonstration of a vari-focal optical see-through head-mounted display using freeform Alvarez lenses. *Opt Express* **27**, 15627–15637 (2019).
11. Huang HK, Hua H. High-performance integral-imaging-based light field augmented reality display using freeform optics. *Opt Express* **26**, 17578–17590 (2018).
12. Hu XD, Hua H. High-resolution optical see-through multi-focal-plane head-mounted display using freeform optics. *Opt Express* **22**, 13896–13903 (2014).
13. Jang C, Bang K, Chae M et al. Waveguide holography for 3D augmented reality glasses. *Nat Commun* **15**, 66 (2024).
14. Blinder D, Birnbaum T, Ito T et al. The state-of-the-art in computer generated holography for 3D display. *Light Adv Manuf* **3**, 572–600 (2022).
15. Fan ZB, Cheng YF, Chen ZM et al. Integral imaging near-eye 3D display using a nanoimprint metalens array. *eLight* **4**, 3 (2024).
16. Liu ZY, Wang DY, Gao H et al. Metasurface-enabled augmented reality display: a review. *Adv Photonics* **5**, 034001 (2023).
17. Kim J, Seong J, Kim W et al. Scalable manufacturing of high-index atomic layer-polymer hybrid metasurfaces for metaphotonics in the visible. *Nat Mater* **22**, 474–481 (2023).
18. Lee GY, Hong JY, Hwang SH et al. Metasurface eyepiece for augmented reality. *Nat Commun* **9**, 4562 (2018).
19. Li ZY, Lin P, Huang YW et al. Meta-optics achieves RGB-achromatic focusing for virtual reality. *Sci Adv* **7**, eabe4458 (2021).
20. Li ZY, Pestourie R, Park JS et al. Inverse design enables large-scale high-performance meta-optics reshaping virtual reality. *Nat Commun* **13**, 2409 (2022).
21. Wang SM, Wu PC, Su VC et al. A broadband achromatic metalens in the visible. *Nat Nanotechnol* **13**, 227–232 (2018).
22. Yao J, Lin R, Chen MK et al. Integrated-resonant metadevices: a review. *Adv Photonics* **5**, 024001 (2023).
23. Liu ZY, Zhang C, Zhu WQ et al. Compact stereo waveguide display based on a unidirectional polarization-multiplexed metagrating in-coupler. *ACS Photonics* **8**, 1112–1119 (2021).
24. Hua JY, Hua EK, Zhou FB et al. Foveated glasses-free 3D display with ultrawide field of view via a large-scale 2D-metagrating complex. *Light Sci Appl* **10**, 213 (2021).
25. Zhang X, Li X, Zhou HQ et al. Multifocal plane display based on dual polarity stereoscopic metasurface. *Adv Funct Mater* **32**, 2209460 (2022).
26. Malek SC, Overvig AC, Alù A et al. Multifunctional resonant wavefront-shaping meta-optics based on multilayer and multi-perturbation nonlocal metasurfaces. *Light Sci Appl* **11**, 246 (2022).
27. Shi YY, Wan CW, Dai CJ et al. Augmented reality enabled by on-chip meta-holography multiplexing. *Laser Photonics Rev* **16**, 2100638 (2022).
28. Tang J, Wan S, Shi YY et al. Dynamic augmented reality display by layer-folded metasurface via electrical-driven liquid crystal. *Adv Opt Mater* **10**, 2200418 (2022).
29. Yi C, Wang ZJ, Shi YY et al. Creating topological exceptional point by on-chip all-dielectric metasurface. *Light Sci Appl* **14**, 262 (2025).
30. Li X, Chen LW, Li Y et al. Multicolor 3D meta-holography by broadband plasmonic modulation. *Sci Adv* **2**, e1601102 (2016).
31. Nikolov DK, Bauer A, Cheng F et al. Metaform optics: bridging nanophotonics and freeform optics. *Sci Adv* **7**, eabe5112 (2021).
32. Wang ZJ, Li Z, Wan CW et al. Gesture-interactive dynamic holo-display via topography flexible metasurfaces. *ACS Nano* **19**, 1286–1294 (2025).
33. Joo WJ, Kyoung J, Esfandyarpour M et al. Metasurface-driven OLED displays beyond 10, 000 pixels per inch. *Science* **370**, 459–463 (2020).
34. Zhao Z, Liu XY, Ji Y et al. Meta-lens digital image correlation. *Opto-Electron Adv* **8**, 250014 (2025).
35. Liu XY, Zhang JC, Leng BR et al. Edge enhanced depth perception with binocular meta-lens. *Opto-Electronic Sci* **3**, 230033 (2024).
36. Luo YX, Li XY, Zhang RZ et al. Monocular metasurface for structured light generation and 3D imaging with a large field-of-view. *ACS Appl Mater Interfaces* **16**, 39906–39916 (2024).
37. Yang R, Wan S, Shi YY et al. Immersive tuning the guided waves for multifunctional on-chip metaoptics. *Laser Photonics Rev* **16**, 2200127 (2022).
38. Wei QS, Huang LL, Li XW et al. Broadband multiplane holography based on plasmonic metasurface. *Adv Opt Mater* **5**, 1700434 (2017).
39. Zhao RZ, Wei QS, Li YZ et al. Stereo Jones matrix holography with longitudinal polarization transformation. *Laser Photonics Rev* **17**, 2200982 (2023).
40. Wang J, Chen J, Yu FL et al. Unlocking ultra-high holographic information capacity through nonorthogonal polarization multiplexing. *Nat Commun* **15**, 6284 (2024).
41. Xiong B, Liu Y, Xu YH et al. Breaking the limitation of polarization multiplexing in optical metasurfaces with engineered noise. *Science* **379**, 294–299 (2023).
42. Sasián J. *Introduction to Aberrations in Optical Imaging Systems* (Cambridge University Press, Cambridge, 2012).
43. Rolland JP, Davies MA, Suleski TJ et al. Freeform optics for imaging. *Optica* **8**, 161–176 (2021).
44. Yang T, Jin GF, Zhu J. Automated design of freeform imaging systems. *Light Sci Appl* **6**, e17081 (2017).
45. Duerr F, Thienpont H. Freeform imaging systems: Fermat's principle unlocks "first time right" design. *Light Sci Appl* **10**, 95 (2021).
46. Bauer A, Schiesser EM, Rolland JP. Starting geometry creation and design method for freeform optics. *Nat Commun* **9**, 1756 (2018).
47. Fuerschbach K, Rolland JP, Thompson KP. Theory of aberration fields for general optical systems with freeform surfaces. *Opt Express* **22**, 26585–26606 (2014).
48. MicroLED 0.13" Display, 2024. https://www.jb-display.com/product_des/1.html.
49. Hsu CW, Zhen B, Stone AD et al. Bound states in the continuum. *Nat Rev Mater* **1**, 16048 (2016).
50. Jia WH, Xu C, Li XL et al. Full-color AR holography by chip-integrated meta-optics. *Nano Lett* **25**, 12269–12275 (2025).

Acknowledgements

The authors acknowledge the funding provided by National Natural Science Foundation of China (U21A20140), National Key Research and Development

Program of China (2021YFA1401200), Beijing Natural Science Foundation (JQ24028), Beijing Nova Program (20240484557), and Synergetic Extreme Condition User Facility (SECUF).

Author contributions

YT.W., T.Y. and L.H. proposed the idea, T.Y., L.G. and J.D. conducted the system design, S.Z., L.H., and Y.Z. conducted metasurface designs and numerical simulations, B.W., and J.L. fabricated the samples, S.Z., L.G., YD.W. and Y.Z. performed the measurements, S.Z., L.G., C.Q., T.Y. and L.H. prepared the manuscript. L.H., T.Y., D.C. and YT.W. supervised the overall projects. All the authors analyzed the data and discussed the results.

Competing interests

The authors declare no competing financial interests.

Supplementary information

Supplementary information is available for this paper at https://doi.org/10.29026/oes.2026.250031

Published by Editorial Office of *Opto-Electronic Science*, Institute of Optics and Electronics, Chinese Academy of Sciences.