Metasurfaces for near-eye display applications
Yan Li, Xiaojin Huang, Shuxin Liu, Haowen Liang, Yuye Ling and Yikai Su

https://doi.org/10.29026/oes.2023.230025

Received: 31 July 2023; Accepted: 7 October 2023; Published online: 31 October 2023

Related articles

Metasurface-based nanoprinting: principle, design and advances
Rao Fu, Kuixian Chen, Zile Li, Shaohua Yu, Guoxing Zheng

Highly efficient vectorial field manipulation using a transmitted tri-layer metasurface in the terahertz band
Huan Zhao, Xinke Wang, Shutian Liu, Yan Zhang

Time-sequential color code division multiplexing holographic display with metasurface
Xin Li, Qinmiao Chen, Xue Zhang, Ruizhe Zhao, Shumin Xiao, Yongtian Wang, Lingling Huang

Integral imaging-based tabletop light field 3D display with large viewing angle
Yan Xing, Xing-Yu Lin, Lin-Bo Zhang, Yun-Peng Xia, Han-Le Zhang, Hong-Yu Cui, Shuang Li, Tong-Yu Wang, Hui Ren, Di Wang, Huan Deng, Qiong-Hua Wang

More related article in Opto-Electron Journals Group website

http://www.oejournal.org/oes
Virtual reality (VR) and augmented reality (AR) are revolutionizing our lives. Near-eye displays are crucial technologies for VR and AR. Despite the rapid advances in near-eye display technologies, there are still challenges such as large field of view, high resolution, high image quality, natural free 3D effect, and compact form factor. Great efforts have been devoted to striking a balance between visual performance and device compactness. While traditional optics are nearing their limitations in addressing these challenges, ultra-thin metasurface optics, with their high light-modulating capabilities, may present a promising solution. In this review, we first introduce VR and AR near-eye displays, and then briefly explain the working principles of light-modulating metasurfaces, review recent developments in metasurface devices geared toward near-eye display applications, delved into several advanced natural 3D near-eye display technologies based on metasurfaces, and finally discuss about the remaining challenges and future perspectives associated with metasurfaces for near-eye display applications.

Keywords: metasurface; near-eye display; virtual reality; augmented reality; 3D display

Near-eye displays are crucial to VR and AR technologies. Innovative near-eye headsets such as Google Glasses, Microsoft Hololens, Meta Quest, and Apple Vision Pro, have been developed and launched, offering unprecedented user experiences that push the boundaries of human imagination, as shown in Fig. 1. However, there are still challenges to overcome before VR and AR can be widely adopted for everyday use. For VR displays, it is critical to provide a large field-of-view (FOV) and high-resolution virtual images to enhance the realism and immersion of the virtual environment. However, this often results in increased bulkiness and weight. For AR displays, compactness and lightweightness are highly desired for long-term viewing and interaction. An adequate FOV, the see-through ability without obvious distortions of the real-world scene, and good contrast in

Introduction

Recently, the emergence of the metaverse has sparked a surge of interest in virtual reality (VR) and augmented reality (AR) technologies, which has led to extensive research in both industry and academia. VR creates fully-immersive virtual environments by presenting digital images to users, while effectively blocking out the real world. And AR allows the viewers to perceive and interact with a mixed-reality world by superimposing virtual information onto real-world scenes. These technologies offer a wide range of potential applications, including entertainment, military training, education, remote communication, and design engineering. It is anticipated that they will soon revolutionize our daily lives, transforming the way we live, work, and learn, and emerging as the next-generation human-machine interface.
both the real-world scene and virtual images are also required. Balancing and optimizing these features simultaneously remains a challenge. Moreover, because of the stereoscopic 3D display technology employed, these near-eye headsets are incapable of generating a realistic 3D effect with all the depth cues correctly presented, as experienced in the real world. This discrepancy in depth cues can result in an unrealistic 3D effect and even cause visual fatigue and nausea, compromising the visual experience for both VR and AR.

Metasurfaces, ultra-thin planar elements consisting of sub-wavelength antennas, offer superior modulation capabilities for light amplitude, phase, and polarization state, outperforming conventional refractive and diffractive optics. Extensive research has been conducted on optical metasurface devices, including gratings, lenses, and holograms. Advances in metasurface device design and fabrication engineering have facilitated their applications in a great variety of areas such as displays, cameras, and microscopes. With their ultra-thin form factor, subwavelength modulation scale, and high modulation flexibility, metasurfaces are promising candidates as key components in near-eye displays to replace bulky conventional optics or enable novel functionalities, paving the way for next-generation AR and VR technologies.

In this review, we first introduce VR and AR near-eye displays, discussing the working principles, basic architectures, key functional components, and challenges. We then briefly explain the working principles of light-modulating metasurfaces. Next, we review recent developments in metasurface devices geared toward near-eye display applications, illustrating their working principles, functionalities, and performance characteristics within various VR and AR architectures. We then delved into several natural 3D near-eye display technologies based on metasurfaces. Finally, we discuss the remaining challenges associated with metasurfaces for near-eye displays and suggest future research directions.

**Principles of metasurfaces**

Metasurfaces, constructed from subwavelength-scaled
meta-atoms arranged in specific configurations, are artificial materials with the ability to independently or simultaneously manipulate various properties of light, including phase, amplitude, wavelength, polarization, and temporal characteristics. As a result, metasurfaces have emerged as an innovative platform for controlling light. Their remarkable functionalities offer significant opportunities for the design of near-eye displays, which necessitate efficient manipulation of diffractive light within a confined space.

In a metasurface, each individual meta-atom can be conceptualized as a secondary wave emitter that generates light with specific optical properties by the Huygens principle. For a near-eye display, the desired capability is to modulate the phase, polarization, and wavelength of light. Therefore, this review will primarily focus on the fundamental principles underlying the modulation of these optical properties.

Phase modulation is one of the most frequently utilized features of metasurfaces in near-eye displays. Various mechanisms, such as resonant phase, propagation phase, geometry phase, and other nonlocal modulations, can be implemented in meta-atoms for phase control.

Within a specific wavelength range, the incident light can excite the resonant mode in meta-atoms, leading to a phase shift in the transmitted or reflected light, known as the resonant phase. This phase is influenced by the geometric parameters and electromagnetic properties of the meta-atoms constituent materials. By tuning these parameters, the resonant condition and resonant phase can be effectively modulated. Metallic meta-atoms are often used for resonant phase modulation due to their ability to utilize plasmonic resonance. However, their practical applications are hindered by the significant ohmic loss, resulting in low working efficiency. As an alternative, dielectric meta-atoms exhibit Mie resonance, offering a more efficient method for resonant phase modulation with minimal ohmic loss. Unlike plasmonic resonance, Mie resonance originates from the interplay between intrinsic electric and magnetic resonances within these meta-atoms. More recently, advanced resonant modes, such as integrated resonance, have been proposed, providing greater degrees of freedom for sophisticated modulation of the resonant phase.

The propagation phase refers to the phase shift that occurs when light traverses a dielectric meta-atom. In this scenario, the meta-atom is treated as a truncated waveguide, where the height and the effective refractive index of the meta-atom determine the phase accumulation during light propagation. Typically, the height is uniform, while the effective refractive index is influenced by the fill factor and constituent material of the meta-atom. Therefore, adjusting the geometrical parameters enables the tuning of the propagation phase. The propagation phase is often utilized in the design of polarization-insensitive metasurfaces.

In contrast to the previous two principles, the geometry phase, also known as the Pancharatnam-Berry phase, provides a unique approach to phase modulation in metasurfaces. When circularly polarized light interacts with an anisotropic meta-atom, the output light undergoes a polarization state flip to orthogonal states, accompanied by an induced phase shift. This geometry phase is precisely twice the in-plane rotational angle of the meta-atoms. Consequently, an entire metasurface can be constructed using an array of these anisotropic meta-atoms, all featuring the same geometry but different in-plane rotational angles.

The aforementioned phase modulation mechanisms primarily focus on local phase modulation. However, nonlocal modulation of phase presents greater opportunities for efficient angular manipulation. This approach relies on both the local Bloch modes and a large number of nonlocal overlapping modes to effectively modulate light. By leveraging these mechanisms, light can interact with the desired diffraction channel with a higher probability, maintaining a very high diffraction efficiency even at large deflection angles. For instance, it has been reported that both transverse electric (TE) and transverse magnetic (TM) diffraction efficiencies can reach approximately 75% at a deflection angle of 75° using a metagrating designed for a wavelength of 1050 nm, based on freeform multimode geometries. Since nonlocal modes are difficult to predict and engineer, these metagratings are typically designed using an inverse design approach, which is result-driven, with the local modes being automatically obtained.

By carefully designing metasurfaces with the phase modulation methods mentioned above, they could achieve superior optical properties such as a wider angular bandwidth and higher diffraction efficiency. Moreover, metasurfaces can modulate not only the phase but also the polarization or amplitude simultaneously. These features allow for convenient polarization multiplexing, providing more design opportunities for light manipulation.
freedoms to improve the optical performance and enable new functionalities in near-eye displays\textsuperscript{79,80}.

**Fundamentals of near-eye displays**

The recent advancements in optics, high-resolution displays, and information technologies have led to the emergence of VR and AR near-eye displays, finding applications in various fields.

A typical VR display architecture is shown in Fig. 2(a). A display panel is placed within one focal-length distance from the eyepiece. This arrangement allows the virtual image of the display panel to be magnified by the eyepiece and viewed by the human eye. With the external light blocked, the viewer becomes fully immersed in the virtual environment. The eyepiece and the image source are two critical components of a VR display. The eyepiece can consist of a group of refractive or diffractive optics, and the system configuration can be either transmissive, as shown in Fig. 2(a), or reflective with a folded optical path for a more compact form factor, as in a pancake scheme\textsuperscript{81,82}. High-resolution liquid crystal displays (LCDs), organic light-emitting-diode (OLED) microdisplays, and micro light-emitting-diode (μLED) microdisplays are commonly employed as image sources\textsuperscript{83} in VR displays.

For VR displays, which aim to provide an immersive virtual environment, a critical challenge is to provide a wide FOV that is comparable to the human visual system. The monocular FOV of the human eye is approximately 160° horizontally and 130° vertically, and the binocular FOV is about 200° horizontally and 130° vertically\textsuperscript{84}. Hence, the eyepiece should have good imaging capability within a large FOV.

VR displays face a significant challenge in the form of the vergence-accommodation conflict (VAC) problem\textsuperscript{28,85,86}, which leads to 3D visual fatigue. This issue arises from the stereoscopic 3D display technique used in commercial VR headsets. When viewing a real-world scene, both eyes converge at the depth of the 3D object, with each eye accommodating (focusing) at the same depth as shown in Fig. 2(b). However, in a stereoscopic VR display, where two separate parallax images are projected to each eye to create a virtual 3D image, the eyes converge on the virtual 3D image, but each eye focuses on the 2D virtual image plane on the display panel, as shown in Fig. 2(c). The depth of the virtual 3D image perceived by two-eye convergence could be varied by adjusting the parallax images, but the depth perceived by eye accommodation remains at a fixed image plane distance when the image distance and eyepiece focal length are set. This mismatch between accommodation and vergence distances leads to visual fatigue, nausea, and other discomfort. Various true 3D display\textsuperscript{87} technologies such as light field display\textsuperscript{80,88−90}, holographic display\textsuperscript{91,92}, and volumetric 3D display\textsuperscript{93,94} have been proposed to address this problem.

For AR displays, there are two basic architectures: video-see-through and optical-see-through\textsuperscript{95}. The former is quite similar to VR, comprising an image source and an eyepiece, but with an additional component - a real-time capturing camera. Instead of viewing the real world optically, the viewer sees the real-time video of the real world captured by the camera, with computer-generated virtual information added to the display, as shown in Fig. 3(e). Recently, Apple Vision Pro, a representative video-see-through AR product, was launched, offering a
high-quality real-world environment with abundant virtual resources. This review, however, focuses on optical-see-through AR displays. An optical see-through AR display requires an additional functional component called an optical combiner, in addition to the display panel and eyepiece. This component overlays the virtual image onto the real world. A combiner typically reflects virtual image light and transmits real-world light into the eye to combine the virtual and the real world. Hence, a beam splitter (BS)\textsuperscript{96,97}, a polarization beam splitter (PBS)\textsuperscript{98,99}, a partially reflective mirror (including half mirror), or gratings\textsuperscript{100} on a waveguide could all serve as a combiner in an optical see-through AR display. Figure 3(a–d) illustrates various optical-see-through AR architectures based on different optical combiners.

Figure 3(a) illustrates the architecture of the simplest optical see-through AR display, which is based on a BS combiner. The virtual image is first magnified by the eyepiece and then reflected towards the eye by the BS. Simultaneously, the real-world light can pass through the BS directly, entering the eye without distortion. The BS could be a regular polarization-independent BS in the form of a cube or thin plate, with its transmittance and reflectance optimized to achieve the desired brightness for virtual and real environments. Obviously, there is a tradeoff between the efficiency of the virtual image and real-world light. However, if the virtual image light is linearly polarized, such as that from an LCD, employing a PBS can significantly improve the efficiency of virtual image light while maintaining ~50% transmittance for the real-world light.

The birdbath architecture is a commonly used design in commercial optical-see-through AR products\textsuperscript{101}. As shown in Fig. 3(b), the image light, initially modulated by the refractive optics near the display panel, undergoes a series of reflections. It is first reflected by the BS, then by the curved partial mirror, and finally passes through the BS again before reaching the eye. The refractive optics and the curved partial mirror work together to magnify the image, functioning as an eyepiece. The folded optical path, combined with the design flexibility offered by the separate refractive and reflective optics, allows these systems to deliver high-quality images and a large FOV. Moreover, they maintain a relatively compact form factor and are lightweight.

Figure 3(c) demonstrates the schematic of an AR display system based on freeform optics. Here, image light injected into the freeform prism is first reflected at the glass-air interface via total internal reflection (TIR), and then by the partially reflective mirror. Each surface of the prism could be individually designed and optimized to jointly achieve high image quality for the virtual image.

| Fig. 3 | Schematics of AR display architectures based on (a) a half mirror/BS combiner, (b) birdbath optics, (c) freeform prisms, and (d) a waveguide with grating couplers. (e) Schematic diagram of a video see-through AR display. |
A compensator prism is required to compensate for the real-world scene. This design can yield high-quality images with a broad FOV and high efficiency. Nonetheless, this approach also results in a bulky and heavy overall system. Furthermore, it often necessitates more complex and costly manufacturing processes.

Figure 3(d) shows the schematic of a waveguide-based AR display utilizing grating couplers. Light from the light engine is first collimated by a collimating lens before being coupled into the waveguide. It then propagates within the waveguide via TIR until it is coupled out towards the eye by the grating. The input coupler is typically highly efficient, maximizing the utilization of light emitted from the microdisplay. In addition to gratings, geometry prisms could also serve as high-efficiency input couplers. Conversely, the output coupler, which also functions as a pupil expander, usually has a lower and spatially varying efficiency to ensure a uniformly duplicated eye box. The out-couplers are usually made of surface relief gratings, volume gratings, or holographic optical elements. Due to the dispersive nature of gratings, waveguide architectures often employ multiple waveguides or a wavelength multiplexing technique to enable full-color rendering. The waveguide architecture provides a smaller form factor and lighter weight compared to other designs. It allows for very high transmittance of most real-world light, but the overall efficiency for the virtual image light is relatively low. Additionally, the FOV of the waveguide-based AR displays is limited due to the constraints imposed by TIR.

In the realm of near-eye display architectures, there is always a trade-off between various features such as FOV, resolution, image quality, and form factor. Researchers and engineers are constantly striving to strike a balance between visual performance and device compactness. While traditional optics are nearing their limitations in addressing these challenges, ultra-thin metasurface optics, with their high light-modulating capabilities, may present a promising solution. In the following sections, we will review recent developments in metasurface devices orientated toward near-eye display applications.

**Metasurfaces in VR displays**

As previously discussed, the architecture of a VR display is relatively straightforward, composed of two main functional components: the image source and the eyepiece. In theory, metasurface devices could serve as either the image source or the eyepiece. However, for VR displays, the image source must provide large-size, video-rate, full-color images to create an immersive virtual environment. This requirement surpasses the capability of the state-of-art metasurfaces. Therefore, the application of metasurfaces in VR displays is primarily restricted to functioning as an eyepiece. Compared to traditional bulky optics, metasurface eyepieces, characterized by their planar architectures with exceptionally compact footprints and versatile phase modulation capabilities, provide greater light deflecting and aberration suppression capabilities. The main challenge for metasurface eyepieces in achieving a full-color VR display with a large FOV lies in the design and fabrication of large-size, achromatic metalenses.

In 2021, Li et al. proposed a method for achieving large-area multiwavelength achromatic metalenses using multiple-zone dispersion engineering, which they subsequently employed in a VR system. While broadband achromatic metalenses could be achieved by designing spatially varied meta-atoms that independently control phase profile and dispersion parameters, such as group delay (GD) and group delay dispersion, the required GD significantly increases with the diameter of metalens, exceeding the capabilities of meta-atoms. This typically restricts the size of achromatic metalenses to tens of microns. To overcome this limitation, they divided the metalens into multiple zones, as shown in Fig. 4(a), to reduce the required GD within each zone and fabricated a 2-mm-diameter, high numerical aperture (NA=0.7) RGB-achromatic metalens. The metalens is composed of TiO₂ nanofin structures on fused silica substrate. Figure 4(b) shows a scanning electron microscopic (SEM) image of the fabricated metalens. Figure 4(c) illustrates the schematic of the VR display system, where the achromatic metalens is used as an eyepiece and an RGB-laser-based image source is employed. The displayed VR images are shown in Fig. 4(d–g). As can be seen, different color components achieve high image quality at the same image depth simultaneously, indicating negligible chromatic aberration.

While the aforementioned method achieves the desired phase profiles, it necessitates a significant variation in meta-atom structures, thereby increasing fabrication complexity. To address this, the same research group proposed a novel inverse-design framework in 2022. This new approach deviates from optimizing the phase profiles and instead focuses on maximizing the intensity of different wavelengths at the focal spot. By using a fast-
approximate solver and an adjoint method, a 1-cm-diameter, RGB-achromatic, polarization-insensitive metalens was designed based on the same materials but simpler meta-atom structures. The picture of the achromatic metalens with an NA of 0.3 is shown in Fig. 5(a). Figure 5(b) shows the measured focal intensity distribution in the XZ plane for the red (488 nm), green (532 nm), and blue (658 nm) colors, respectively, demonstrating minimal focal shift. The focusing efficiency of the fabricated metalens was measured to be 15.8%, 13.6%, and 16.1% for RGB colors, respectively. Figure 5(c) shows a photo of the VR system employing the achromatic metalens eyepiece and a laser-illuminated micro-LCD, while Fig. 5(d–g) exhibits the VR images generated by the system.

Despite these advances in large-scale metalens design engineering, the fabrication of metalenses still heavily relies on high-resolution nanopatterning techniques like electron-beam lithography, which are usually high-cost and low-throughput. Recently, Kim et al. proposed and demonstrated a cost-effective, high-throughput method for mass production of large-aperture visible metalenses using deep-ultraviolet argon fluoride immersion lithography and wafer-scale nanoimprint lithography. The schematic of the mass production procedures is shown in Fig. 6. Using this novel method, they fabricated hundreds of 1-cm-diameter metalens via a 12-inch master stamp. A prototype of the VR device was demonstrated with a high-performance, mass-manufactured metalens capable of displaying red, green, and blue color images.

**Metasurfaces in AR displays**

Given their versatile functionality, high optical performance and ultra-thin form factors, metasurfaces have been proposed as critical optical components in various AR display architectures. They can function as eyepieces, combiners, and more, substituting traditional bulky optics or integrating new advanced optical functionalities, to achieve more compact and lightweight, high image-quality and large FOV AR displays. Though not as sophisticated as video-rate refreshing microdisplays, metasurfaces have also been employed in some AR displays as an image source to provide holographic images. In the following sections, we will review the applications of metasurfaces in different AR architectures.

**Base on beam splitters**

As discussed before, AR displays based on beam splitters typically involve a transmissive on-axis refractive lens as an eyepiece and a beam splitter, occasionally supplemented by a partial reflective lens, as the combiner. Within such architectures, transmissive metalenses emerge as a compelling alternative to conventional refractive lenses, offering the potential for a more compact form factor and an expanded FOV.

In 2018, Lee et al. demonstrated an AR display featuring a 90° FOV, facilitated by a 20-mm large-area transmissive metasurface eyepiece within a beam-splitter-based architecture. The large-area metalens, composed of poly-crystalline silicon posts on a SiO$_2$ substrate, gets...
its phase modulated by the geometry phase principle. The schematic of the AR display is shown in Fig. 7(a). The image light from the display panel first traverses the BS, after which its different color components are reflected by the corresponding dichroic mirrors. Upon its second encounter with the BS, the image light is reflected and subsequently modulated by the metalens, forming a magnified virtual image in the viewer’s line of sight. The circular polarizer arrangement in Fig. 7(a) allows the metalens to function concurrently as a positive lens for.
the image light and as a transparent window for real-world light. Hence, the viewer perceives the virtual image floating on the real world. In this configuration, the dichroic mirrors are positioned at different locations to counteract the chromatic dispersion of the metalens, thereby ensuring that different color images are formed at the same depth. The metalens is fabricated by electron-beam lithography and nanoimprint technique, and its SEM photo is depicted in Fig. 7(b). The resultant image, shown in Fig. 7(c), reveals that the real-world scene remains visible, albeit with decreased brightness due to the use of the circular polarizers, while the RGB virtual image spans a large FOV.

As a matter of fact, the metalenses utilized in VR displays, as we discussed in the previous session, could also be used as an eyepiece in the BS-based AR displays, given their inherent functionality. For instance, Li et al. applied the RGB achromatic metalens, previously introduced in the context of VR display, in a full-color AR display prototype\textsuperscript{102}. Owing to the achromatic nature of the metalens itself, the optical configuration is significantly simplified. As shown in Fig. 7(d), the color image is magnified by the metalens and subsequently reflected towards the eye by the beam splitter. Figure 7(e–f) shows the AR images generated by the prototype, demonstrating that the real-world scene retains a reasonable level of brightness, and the RGB virtual images appear at the same depth, a testament to the achromatic properties of the metalens.

**Base on waveguides**

In waveguide architecture, the key functional components are the collimating lens and the in- and out-couplers. Metasurfaces, with their versatile optical functionality, could act as either a collimating lens or a coupler\textsuperscript{80,108–111}. However, most research focuses on the latter, fully leveraging the high flexibility in wavefront control\textsuperscript{112}. The primary challenges lie in achieving a large FOV and ensuring large and uniform brightness at the eye box.

In 2018, Shi et al. proposed the use of polarization-dependent metagratings to achieve wide FOV waveguide displays\textsuperscript{113}. In such a system, both the in- and out-couplers are polarization-selective metagratings, achieved by spatially interleaving TE-pass and TM-pass gratings. These gratings would diffract TE and TM light from the same incident angle to different output angles. As shown in Fig. 8(a), light from the left and right halves of the FOV is encoded with TE and TM polarizations, respectively. Each polarization is guided into and out of the waveguide. In this design, a specific guided angle is mapped to two different output angles. Although the
The maximum FOV for each polarization is still limited by the TIR condition, the separation of the TE and TM FOVs via the polarization-selective meta-couplers significantly increases the total FOV. The theoretical calculation indicates that the proposed design can achieve a FOV 70% larger than the maximum horizontal FOV achieved in conventional diffractive grating designs.

In 2022, Boo et al. proposed and demonstrated a novel large-FOV, full-color AR display prototype based on a single layer of metasurface waveguide. The schematic of the prototype is shown in Fig. 8(b). A laser scanner projects an image with a minimal divergence angle and a small FOV. The in-metasurface optical element (MOE) then diffracts the image light, guiding it along the waveguide in compliance with the TIR condition. Because of the narrow FOV, the light beam retains a nearly identical optical path inside the waveguide. The out-MOE subsequently deflects and converges the light from all FOV angles towards the observer’s eye with relatively strong optical power. Therefore, the eye would observe the image with a large FOV. As the metasurface optical elements are optimized for low dispersion, the same waveguide can facilitate RGB color laser light, resulting in a simplified structure. Fabrication of the MOEs involves a silicon foundry process, incorporating nitride film deposition, deep-ultraviolet lithography, and MOE nano-patterning. In this work, they employed continuous irregular structures other than discrete structures, as illustrated in the SEM image of a fabricated MOE in Fig. 8(c). The preference for continuous structures holds significant promise in terms of development prospects. It circumvents issues associated with phase under-sampling and resonance between adjacent structures, thus leading to enhanced efficiency and broader bandwidth, which are critical factors for advancing metasurface-based AR displays. And Fig. 8(d) shows the displayed AR

---

**Fig. 8** (a) Schematic of a polarization-dependent metagrating-based out-coupler. (b) Schematic of a large-FOV, full-color AR display prototype based on a single-layer metasurface optical element, (c) SEM image of an optimized full-color metasurface element, and (d) AR display result. (e) Schematic of an on-chip metasystem for AR display. (f) Schematic of the inverse-designed metagrating architecture with wavelength-demultiplexing functionality. (g) Simulated in-coupling efficiency of the metagrating at opposite ports in the visible regime. (h) AR display results. The scale bar is 5 μm in (c), and the inset in (d) is the original image. Figure reproduced with permission from: (b–d) ref. 114, Nature Publishing Group, under a Creative Commons Attribution 4.0 International License; (e–h) ref. 119, American Chemical Society.
image captured by a camera. The importance of a large and uniform eye box in waveguide AR displays cannot be overstated, as it ensures that no portion of the image is lost or perceptible brightness variation is observed as the viewer’s eye moves. Long et al. optimized metagratings by using a physics-driven neural network to achieve a high uniformity pupil expansion in a color waveguide display. This optimization required different efficiencies of the out-coupling metagratings at various positions among multiple output channels to ensure uniform output light. Rigorous coupled-wave analysis served to calculate the forward and adjoint electromagnetic field based on the structural parameter. To prevent cross-talk, they used two waveguides: one for blue and green colors, and the other for red. The former involves achromatic metagrating.

More recently, an intriguing on-chip AR waveguide display was introduced. This design features a polarization selective metagrating that functions as the in-coupler, and two metasurfaces simultaneously functioning as the out-coupler and the image source. As shown in Fig. 8(e), free space red and green laser light is coupled in by a wavelength-selective metagrating, then propagates in opposite directions within the waveguide. Each light wave then interacts with a metasurface hologram out-coupler, which deflects it back to free space to generate a holographic image. Here, the in-coupler metagrating is made of amorphous silicon (α-Si) on a planar Si$_3$N$_4$ waveguide, with a thick SiO$_2$ substrate (~500 μm) underneath serving as the bottom cladding, as shown in Fig. 8(f). The calculated in-coupling efficiency of wavelength demultiplexing in the visible regime is plotted in Fig. 8(g). It indicates strong wavelength-selective contrasts of coupling efficiencies for leftward and rightward waveguide propagation at the green (550 nm) and red (650 nm) wavelengths. The experimental result is shown in Fig. 8(h), where a green apple is displayed near the left metasurface out-coupler and red cherries near the right one.

**Base on direct projection**

In the direct-projection architecture of AR systems, the most critical optical element is the see-through reflective lens. This component not only converges reflective virtual image light but also transmits real-world light. Consequently, it combines the functionality of an eyepiece and a combiner in one device, fostering a more compact and highly integrated AR display system. Despite the significant versatility of metasurfaces, traditional metasurfaces are either transmissive or reflective. Achieving such unique new optical features requires great efforts in metasurfaces engineering. Nevertheless, in recent years, a few research groups have ventured into pioneering the development of such innovative metasurface devices. Their work promises to contribute significantly to AR displays, paving the way for more compact, lightweight, and high-performance next-generation AR displays.

The concept of a reflective metasurface visor for near-eye displays could be dated back to 2017. A flat metasurface based on cylindrical silicon pillars was designed to achieve an imaging effect akin to a freeform mirror. By adjusting the duty cycle of the unit structure, the phase profile of the metasurface could be precisely controlled. They have achieved a near-eye display design with a 70° FOV and a good modulation transfer function (MTF) curve, as confirmed by Zemax simulation. However, in their design, the silicon-based metasurface has not been optimized for high transparency to enable the see-through property. Later, in 2021, Avayu et al. designed a full-color, metasurface-based AR visor using three metallic metasurface layers. They fabricated proof-of-concept small-scale samples with two individual metasurface layers. Preliminary test results reveal that green light can be reflected with its wavefront modulated, although the see-through property of the metasurface visor was not thoroughly investigated. Bayati et al. reported simulation work on the design of an achromatic see-through AR visor based on composite metasurfaces with a large FOV (>77°) both horizontally and vertically, and good (>70% transmission and no distortion) see-through quality.

In 2022, Li et al. experimentally demonstrated a planar multifunctional dielectric metasurface visor that could realize the focusing effect for oblique incident red light, while maintaining good see-through property for normal incident light. The optical function of the metasurface visor is illustrated in Fig. 9(a). Thus, the multifunctional visor can simultaneously perform as an eyepiece for the oblique virtual image light and as a combiner that reflects the virtual image light into the viewer’s eye while transmitting light from the real environment. As a result, the whole system could be substantially simplified and compacted, as shown in Fig. 9(b). These unique optical properties are obtained by using geometry phase modulation via the spatial orientation variation of the rectangular silicon nanostructures with a high aspect ratio on a transparent sapphire substrate.
The SEM image of the fabricated metasurface is shown in Fig. 9(e). The experimental result confirms an effective focusing effect, as shown in Fig. 9(c), and spectrum measurements, shown in Fig. 9(d), indicate reasonably high transparency, which is consistent with the simulation. Based on the metasurface visor, the researchers further developed an AR display prototype, rendering RGB virtual letters on the real world without needing an additional beam splitter. The display result is shown in Fig. 9(f, g) where the real-world scene is observed directly through the metasurface visor. This marks the first experimental demonstration of AR images generated by a single-piece metasurface visor.

Despite the promise of geometry-phase-based metasurface visors, they are, however, chromatic and polarization-dependent. Recently, Luo et al. proposed an RGB achromatic and polarization-insensitive trans-reflective single-layer metalens for AR displays. They constructed a structural phase library that incorporated silicon meta-units of various shapes and dimensions, as shown in Fig. 10(a), to achieve high reflectance and cover a wide range of phase responses. From this library, an optimized set of structures was selected and positioned across the metasurface so that RGB colors can accomplish the desired phase profiles simultaneously, as shown in Fig. 10(b). This arrangement effectively facilitates achromatic focusing. Moreover, because of the center-symmetry of the employed meta-units, the metalens exhibits sensitivity to the polarization states of the incident light. Fig. 10(c, d) shows the simulated focal spots of red, green, and blue under different linear-polarized collimated incident light. Importantly, all colors converge at the same focal point, and the focusing behaviors remain consistent across two polarization states.

In the abovementioned metasurfaces, localized modes predominate, wherein individual meta-units govern the wavefront shape over a broad bandwidth, thereby linking the phase properties of neighboring wavelengths closely. Nonlocal lattice modes, instead, extend over many unit cells to support high resonant diffraction efficiency at a specific wavelength, while maintaining low efficiency for most other wavelengths. The resonance properties inherent in this approach could be harnessed to develop an RGB achromatic see-through metasurface visor for AR displays.

Malek et al. put forward a concept for multilayer and multi-perturbation nonlocal dielectric metasurface systems based on spatially varying geometry phase. This design allows for the multiplexing of independent wavefront-modulating functionalities across multiple resonant wavelengths while maintaining transparency over the rest of the spectrum. They suggested the use of the non-local metasurface system as an optical see-through lens in an AR display, capable of reflecting selected narrowband wavelengths of contextual information to the viewer’s eye, while simultaneously providing an unobstructed,
broadband view of the real world, as schematically shown in Fig. 11(a). The structure of the metasurface system is shown in Fig. 11(b). It involves a single-function metasurface on top of the substrate based on $p_2$ meta-units operative at the green wavelength, and at the bottom, a dual-function metasurface based on $p_1$ meta-units operative at the red and blue wavelengths. Both metasurfaces are composed of rectangular apertures
etched into a thin film of TiO\textsubscript{2} coated with an antireflection layer of SiO\textsubscript{2}. The simulated transmission and reflection spectra of the metasurface system are shown in Fig. 11(c), indicating high broadband transmission of real-world light and a narrowband reflection at the three chosen visible wavelengths. As such, these novel nonlocal metasurface systems may serve as a strong contender for a multifunctional, ultrathin optical element in AR displays.

The development of high-performance reflective metasurface visors has recently broadened its scope, moving beyond planar structure to explore freeform curved substrates. Nikolov et al. introduced a versatile design-to-fabrication process for creating metasurfaces on a freeform substrate, the metaforms, to combine the advantages of freeform optics and metasurfaces\textsuperscript{125}. Such a metaform can be integrated into an AR display as a combiner as sketched in Fig. 11(d). A miniature 1.5 mm by 2 mm imager based on a metaform mirror was demonstrated with a visible wavelength of 632.8 nm, as shown in Fig. 11(e). The Ag-SiO\textsubscript{2}-Ag metasurface is formed on the concave brass substrate. Employing layers of metallic materials, the freeform metasurface device effectively performs the reflective lens function over varying spatial frequencies in object space of up to 10.42 lp/mm, as shown in Fig. 11(f). However, the see-through functionality needs to be further investigated by adding see-through holes in the metals or optimizing the design with dielectric materials.

**Metasurfaces-based natural 3D near-eye displays**

Metasurfaces, as we have discussed, hold great potential for applications in VR and AR near-eye displays to improve imaging performance, enlarge FOV or increase compactness. However, most near-eye displays still employ stereoscopic 3D display technology, which can lead to VAC and visual fatigue. To address this issue, natural 3D display approaches have been proposed, including Maxwellian viewing display, holographic display, light field display, and multi-/vary-focal displays. Thanks to the versatility of metasurfaces, metasurface devices have been employed as important optical elements in these advanced 3D near-eye displays to help alleviate visual fatigue and provide a more natural 3D visual experience\textsuperscript{126}.

Recently, Song et al. proposed a Maxwellian-viewing near-eye display that can provide accommodation-free virtual images by using a small aperture (360 μm × 360 μm) transparent Huygens’ metasurface hologram as the display device\textsuperscript{127}. As shown in Fig. 12(a, b), the metasurface, when illuminated by a laser source, generates a holographic image that is then projected to a minuscule point at the pupil position by the eyepiece. In such a pinhole-like Maxwellian-viewing display system\textsuperscript{128}, the image received on the retina remains consistently sharp and clear, irrespective of any changes in the focus accommodation of the crystalline lens. Hence, depth
information cannot be inferred from eye accommodation, and depth perception is solely determined by the vergence of both eyes. This accommodation-free display, therefore, does not induce depth mismatch or 3D visual fatigue. The compact dimension of the metasurface hologram allows for a shorter optical path length, enabling the creation of a lightweight, compact wearable prototype weighing only ~50 grams, as shown in Fig. 12(c). Figure 12(d, e) displays an AR image produced by the prototype. The virtual image maintains its sharpness as the camera adjusts its focus from near to far, thereby revealing its accommodation-free property. However, the Maxwellian-viewing displays are limited by a small eye box, which would lead to image loss when the eyeball moves around. Pupil duplication and multiview display techniques\textsuperscript{122} based on the metasurfaces have been proposed to solve this problem.

Wang et al. demonstrated a near-eye display system that combines a 5-mm-diameter metalens eyepiece with 3D computer-generated holography (CGH)\textsuperscript{129}. Holography is often considered the ultimate display technology as it can provide comprehensive depth information of 3D objects. The combination of metalens and CGH offers a solution for compact lightweight natural 3D near-eye displays. The metalens eyepiece, composed of silicon nitride anisotropic nanofins, is fabricated with a focal length of 6 mm (\(\text{NA}=0.4\)), and a diffraction efficiency of 15.7\% for 532 nm. Based on the metalens, they implemented a VR prototype that projects holographic 3D images to the viewer with an eye relief of 9 mm and an FOV of about 31\°, providing full monocular focus cues without VAC and ensuring a realistic viewing experience for the viewer shown in Fig. 13(a). Figure 13(b, c) shows the photo and the phase distribution of the metalens eyepiece, respectively. As shown in Fig. 13(d–g) a 3D virtual scene consisting of multiple virtual images “Z” “J” and “U” at different depths is rendered on the real-world scene.

Integral imaging display is a type of light field display that could provide VAC-free 3D images based on a high-resolution display screen and a microlens array as shown in Fig. 14(a). Fan et al. proposed the use of a broadband achromatic metalens array to realize integral imaging display\textsuperscript{130}. Figure 14(b–d) shows the SEM photos and the optical photo of the metalens array. The fabricated microlens array consists of 60 × 60 metalenses made of silicon nitride nanostructures on a silicon dioxide substrate. These achromatic metalenses, designed based on effective refractive index theory, each possess a diameter of 14 \(\mu\text{m}\) and an average focal length of 81.5 \(\mu\text{m}\) (from 430 nm to 780 nm). Implemented using the optical setup in Fig. 14(e), the integral imaging display system employing the achromatic microlens array is capable of generating virtual images “3” and “D” at the same or different depths as shown in Fig. 14(f–h). The in-focus/out-of-focus effect is clearly discernible when the camera adjusts its focus, indicating the presence of correct accommodation depth cue. Moreover, because of the achromatic feature of the metalens array, the depths of the different color components remain consistent.

---

**Fig. 13 |** (a) Schematic of a holographic near-eye display based on a metalens eyepiece. (b) Photo of the Metalens eyepiece. (c) Phase distribution of the metalens eyepiece. (d) Original layered model of letters “ZJU”. (e-f) Reconstructed virtual images when the camera is focused on “Z” “J” and “U”, respectively. Figure reproduced with permission from ref.\textsuperscript{129}, MDPI, under an open-access Creative Common CC BY license.
this metasurface-based display system is not specifically intended for near-eye applications, the compact nature of the metalens array allows for the potential design of portable and wearable integral imaging displays for VR and AR applications.

The vari-/multi focal display\textsuperscript{131−133} represents another 3D display technique capable of addressing the VAC issue and providing a natural 3D effect. This is achieved by reconstructing a 3D scene through the display of its 2D slices at different depths. For near-eye vari-/multi-focal displays, one effective approach for rendering virtual images at different depths involves altering the focal length of the eyepiece. Apparently, in vari-/multi-focal near-eye displays, the varifocal eyepiece lens is the most critical electro-optical component. Various varifocal optical elements have been suggested to fulfill this function, such as liquid lens\textsuperscript{131}, liquid crystal (LC) lens\textsuperscript{134}, and deformable mirror\textsuperscript{135}. Recently, Zhu et al. proposed an electrically controllable varifocal metalens operating in the near IR region with the assistance of an addressed nematic LC layer\textsuperscript{136}. By manipulating the voltage applied to the LC, the focal length of the metalens could be adjusted. However, the fabrication of a considerable number of super-pixel cells for this metalens is still challenging. If further optimized for operation in the visible range, such varifocal metalenses could pave the way for high-performance vari-/multi-focal near-eye displays.

**Discussion**

Ultra-thin metasurface devices exhibit great potential for realizing compact and lightweight near-eye displays. In addition to providing viable alternatives to traditional bulky optical components, metasurfaces open the door to unique optical functionalities previously unattainable with conventional devices, due to their exceptional light modulation capabilities and high design freedom. For example, achromatic metasurface devices can compensate color aberration in an ultra-thin form, which is highly attractive for full-color near-eye displays; the nanoscale meta-atoms could enable large diffraction angle, which is crucial for achieving large FOV; moreover, they could even manifest different optical behaviors for transmitted and reflected waves, or multiplex functionalities. Nonetheless, to truly harness the vast capabilities of metasurfaces in near-eye displays and facilitate practical VR/AR applications, several challenges must be addressed.

One significant challenge is the design and fabrication of high-efficient and large-size metasurfaces. Currently, achieving high efficiency in metasurface devices, such as
achromatic metalenses or partial-reflective metasurface devices, is still difficult. The low focusing efficiency can lead to contrast reduction of virtual images and limit the practical application scenarios of AR displays. Though relatively-large-aperture 10 mm metalens eyepieces have been demonstrated, however, to provide large FOVs, the aperture of metasurfaces should be made even larger. And that requires more complicated variation of metaatom structures, posing hurdles from both design and fabrication perspectives. Mass production of metasurfaces, especially those working at visible wavelengths, requires productive, high accuracy, and large area fabrication techniques. Extreme ultraviolet or deep-ultraviolet lithography\textsuperscript{137} steppers are highly recommended for achieving these goals. Exploring alternative fabrication methods such as nano-imprint\textsuperscript{138}, self-assembly\textsuperscript{139}, and laser writing\textsuperscript{140,141} will enrich the variety of fabrication options.

Another limitation is the fixed functionality of metasurface optical devices once they are fabricated. Tunable metasurface devices\textsuperscript{142–144} are highly desirable as they can be employed for VR/AR display applications in terms of metasurface holographic image sources\textsuperscript{43,145–148}, vari-focal metalenses\textsuperscript{149}, and so on. Though some metasurface holograms have been suggested to provide images for the near-eye displays, the generated images are mostly static or could only be switched among a few different states. For good VR or AR experience, the image source should provide full-color high-resolution virtual images at video rate (>24 Hz) to enable dynamic content rendering without flickering. To date, achieving dynamically tunable metasurfaces remains a challenge, primarily due to constraints related to the optical properties of the metasurface materials. The introduction of liquid crystal materials\textsuperscript{150–152} or other tunable materials\textsuperscript{153,154} to metasurface systems shows promise in improving tunability. However, the tuning range, resolution, grey level counts, and response time are still limited. Exploring how to develop tunable metasurfaces with large tuning ranges, high precision, fast response times, and pixel-level addressability is a worthwhile avenue for future research.

From a system perspective, simply replacing some bulky refractive optics with ultra-thin metasurfaces does not necessarily guarantee a more compact form factor for the entire system. The compactness and weight reduction in near-eye display systems are not solely determined by the thinnest devices, but more by the system configuration. Other devices that cannot be replaced by metasurfaces and the need for air or medium-filled spaces between different devices to enable certain optical functionalities can still limit the compactness and weight reduction in near-eye display systems. Moreover, to overcome the VAC problem and reduce visual fatigue, various natural 3D display techniques have been proposed for near-eye displays. However, these techniques increase system complexity and demand more new functional optical components. Effective improvement in the compactness, weight, and performance of near-eye display systems, fully utilizing the advantages of metasurfaces, requires the cooperation of researchers from both metasurface and display areas.

Despite these challenges, we firmly believe that metasurfaces hold great potential for the next generation of AR/VR near-eye displays. Continued research and development efforts in addressing these challenges will contribute to unlocking the full capabilities of metasurfaces for enhanced near-eye display experiences.

**Conclusion**

In this review, we have provided an extensive overview of the recent advancements in the applications of metasurfaces for near-eye displays. We introduce the fundamentals of near-eye display technologies and metasurfaces, review the recent developments of metasurface devices employed in different VR and AR architectures, and then delve into a few metasurface-based natural 3D near-eye display technologies aiming to provide more realistic and comfortable viewing experience. Thanks to their ultra-thin form factors and exceptional flexibility in light modulation, metasurfaces offer a promising solution for the development of compact and lightweight VR/AR headsets. However, several challenges need to be addressed to fully realize the potential of metasurfaces in near-eye display applications. Continued research and development efforts in addressing these challenges will contribute to unlocking the full capabilities of metasurfaces for enhanced near-eye display experiences, and we firmly believe metasurfaces hold great potential for the next generation of VR/AR near-eye displays.

**References**


46. Wang Z, Zhang HR, Hao HT, Cui TJ, Li LL. Intelligent electro-
magnetic metasurface camera: system design and experimental results. *Optics Express* 20, 3822–3828 (2012).
85. Wang YJ, Lin YH. Liquid crystal technology for vergence-ac-


The authors declare no competing financial interests.

Acknowledgements

We are grateful for financial supports from the National Key Research and Development Program of China (2021YFB2802100) and the National Key Research and Development Program of China (62075127 and 62105203).

Author contributions

Y Li framed and wrote most of the manuscript. XI Huang drew and arranged the figures, obtained the copyrights, and provided the technical details of the reviewed papers. SX Liu drew the figures and wrote part of the manuscript. HW Liang wrote part of the manuscript. YY Ling polished the English language. YK Su determined the topic and content.

Competing interests

The authors declare no competing financial interests.