Review Article

Metasurface-based computational imaging: a review

Xuemei Hu⊙**,**[a,b] **Weizhu Xu,**[a,b] **Qingbin Fan,**[a,b] **Tao Yue,**[a,b] **Feng Yan,**[a,b] **Yanqing Lu,**[a,c] **and Ting Xu**[a,c,*]

[a]Nanjing University, Collaborative Innovation Center of Advanced Microstructures, National Laboratory of Solid-State Microstructures, Nanjing, China
[b]Nanjing University, School of Electronic Sciences and Engineering, Nanjing, China
[c]Nanjing University, College of Engineering and Applied Sciences, Jiangsu Key Laboratory of Artificial Functional Materials, Nanjing, China

Abstract. Metasurface-based imaging has attracted considerable attention owing to its compactness, multifunctionality, and subwavelength coding capability. With the integration of computational imaging techniques, researchers have actively explored the extended capabilities of metasurfaces, enabling a wide range of imaging methods. We present an overview of the recent progress in metasurface-based imaging techniques, focusing on the perspective of computational imaging. Specifically, we categorize and review existing metasurface-based imaging into three main groups, including (i) conventional metasurface design employing canonical methods, (ii) computation introduced independently in either the imaging process or postprocessing, and (iii) an end-to-end computation-optimized imaging system based upon metasurfaces. We highlight the advantages and challenges associated with each computational metasurface-based imaging technique and discuss the potential and future prospects of the computational boosted metaimager.

1 Introduction

Metasurfaces are a type of artificial, two-dimensional (2D) material composed of subwavelength nanostructures arranged in a specific pattern. The unique arrangement and design of these nanostructures allow metasurfaces to interact with light and other electromagnetic waves in a highly controlled and tailored manner. The modulating capability of metasurfaces has enabled diverse imaging techniques, such as microscopic imaging,[1] hyperspectral imaging,[2,3] full-Stokes polarization imaging,[4,5] and full-space 3D imaging.[6] This showcases its immense potential in various application fields, including microscopy, spectroscopy, depth sensing, machine vision, and other imaging scenarios. However, the development of metasurface-based imaging techniques faces two main challenges that warrant in-depth research. First, nonidealities are introduced by metasurfaces, such as strong dispersion effects over different wavelengths and fabrication quality. Second, although metasurfaces offer flexible multidimensional design capabilities, how

to construct a general and scalable imaging framework that harnesses and fully exploits the multidimensional modulation capability of metasurfaces—to enhance or revolutionize existing imaging systems—remains an open question.

Computational imaging, situated at the intersection of optics, electronics, signal processing, and machine learning, aims to incorporate computational techniques into the imaging and reconstruction processes. By embracing computation, not only does it significantly relax the constraints of optical system design, but it also opens up new possibilities for expanding imaging capabilities. The introduction of computation into illumination, optics, sensing, and processing greatly extends observation capabilities, enabling the capture of various dimensions of the plenoptic light field. Examples of these capabilities include superresolution microscopic imaging,[7,8] wide field-of-view (FoV) imaging,[9] lensless imaging,[10,11] and non-line-of-sight imaging.[12,13] Traditionally, different light modulators, such as film masks, digital micromirror devices (DMDs), spatial light modulators (SLMs), and diffractive optical elements (DOEs) have been employed to introduce computation in imaging systems. However, these approaches often come with significant system complexity or volume, and their

*Address all correspondence to Ting Xu, xuting@nju.edu.cn

modulation flexibility is limited. As a result, there is a strong demand for flexible and highly integrated modulation techniques in computational imaging to fully exploit the power of computation for designing imaging systems.

In light of recent advancements in metasurface and computational imaging techniques, computational strategies have been integrated into metasurface-based imaging, significantly enhancing the capabilities of current imaging systems. In this paper, we present a comprehensive review of metasurface-based imaging methodologies from the perspective of computational imaging. Specifically, we first review the existing metasurface-based computational imaging from the perspective of plenoptic dimension modulation, i.e., spectrum, polarization, phase, and compound modulation. We then proceed to categorize the existing metasurface-based computational imaging frameworks based on where computation is introduced. This involves computational illumination, computational sensing, computational reconstruction, and a detailed discussion of how computation contributes to improved imaging performance. To conclude the review, we address the primary challenges encountered in building metasurface-based computational imaging systems, along with a forward-looking discussion. Our intention is to provide valuable insights and identify the primary obstacles faced by researchers in these two domains, with the hope of fostering further advancements in computational imaging and nanophotonics.

A brief overview of the proposed paper is included in Fig. 1. As shown in Sec. 2, we first provide a brief overview of the development of computational imaging, discussing the potential advantages of metasurfaces over traditional computational imaging techniques. In Sec. 3, we review works on computational imaging based on metasurfaces, starting from the modulation of spectrum, polarization, angle, depth, and compound dimensions. In Sec. 4, we outline the computational photography framework of existing works in Sec. 3, from the hardware imaging system to computational algorithms. Finally, we discuss the current challenges of metasurface-based imaging from four different perspectives, pointing out possible future research directions and perspectives.

2 Brief Overview of Computational Imaging

The evolution of imaging technology, stretching from the rudimentary principles of the ancient camera obscura to the sophisticated domains of digital photography, represents a remarkable journey in visual documentation. This progression began with the discovery of pinhole imaging, an ancient technique in which a small aperture in a darkened space projected an inverted image onto a surface. This fundamental imaging principle laid the groundwork for understanding how light could be manipulated to capture images. A significant milestone was achieved in the 1830s with the invention of the daguerreotype by Louis Daguerre and the concurrent development of the calotype process by William Henry Fox Talbot, heralding the inception of photography. These methods, employing varied techniques to capture and fix an image onto a surface, revolutionized the methods of image recording and representation. The most transformative development in recent photographic history began in the 1970s and 1980s with the advent of the charge-coupled device. Digital cameras, using electronic sensors to capture images and convert them into digital data, started replacing traditional film cameras. This innovation facilitated easier storage, manipulation, and sharing of images, significantly broadening the applications and accessibility of photography.

Computational imaging has seen significant development since the early 1990s. Unlike the traditional one-to-one direct mapping and capturing of target scene information, computational imaging proposes to introduce computation to both the physical process of capturing images and their subsequent reconstruction. This approach fundamentally revolutionizes the concept of imaging, moving beyond the need for a one-to-one recording of scene points. By integrating computational techniques, more complex and versatile image processing and reconstruction can be enabled, significantly expanding the capabilities and applications of imaging technology. Over decades, computational imaging techniques have significantly evolved, finding applications across various domains. These techniques preliminarily aim to enhance imaging performance, focusing on methods that achieve higher spatial resolution,[14] imaging speed,[15] dynamic range,[16] and multidimensional imaging capabilities (such as depth[17] and hyperspectral imaging[18]). Beyond improving certain aspects of imaging performance, computational imaging has also been designed to realize previously unattainable imaging capabilities. Examples include capturing black hole images,[19] which were once considered impossible, achieving superresolution that surpasses the diffraction limit,[8] light-in-flight imaging that visualizes the movement of light,[20] gigapixel imaging,[21] etc. These advancements illustrate the remarkable progress and versatility of computational imaging technologies.[22]

To achieve the above goals, computational imaging systems commonly require various light modulation elements that manipulate the properties of light, e.g., intensity, spectrum, polarization, and phases, with delicately designed coding schemes so that the required information of light can be efficiently captured. After that, the computational algorithms with
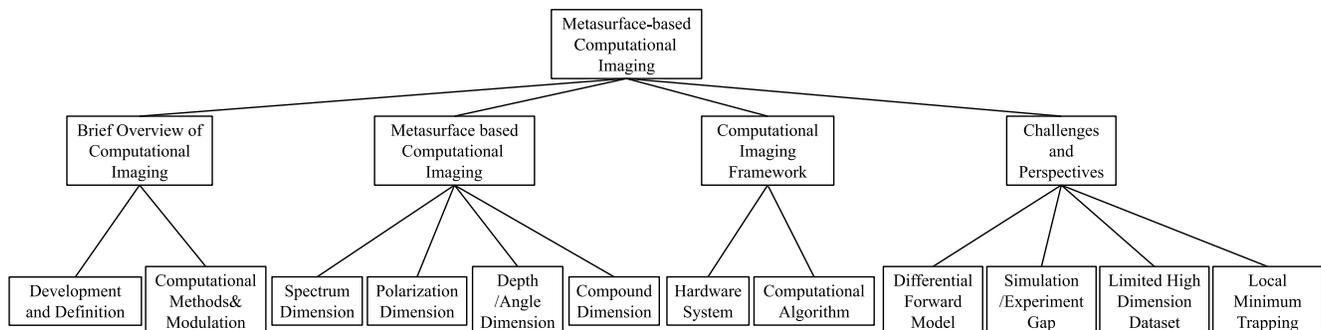

Fig. 1 Brief overview of the structure of the review.

either the conventional iterative optimization algorithms or the neural network-based algorithms are applied to reconstruct the latent images from the measurements. Traditionally, the modulation function of computational imaging systems is accomplished by optical path modulation modules composed of various traditional optical elements, such as light SLMs, DMDs, prisms, beam splitters, lenses, and polarizers. These systems, however, usually face challenges like complex optical paths, bulky system volume, aberrations difficult to control, and restricted modulation capabilities, which have severely hindered the development and application of computational imaging techniques. As an advanced solution, metasurfaces, which are composed of nanostructures on a subwavelength scale, have been introduced, providing multidimensional optical control in an exceptionally compact format. These advanced materials facilitate precise light-wave manipulation, enabling changes in phase, amplitude, spectrum, and polarization while maintaining system compactness and simplicity. The adoption of metasurfaces also leads to a reduction in the number and complexity of components in the optical path, enabling compact, adaptable, and superior-quality computational imaging systems.

Additionally, from a whole system standpoint, the evolution of computational imaging techniques has been remarkable,[23–25] transitioning from the individual design of acquisition systems and the reconstruction algorithm to the end-to-end designing framework, which jointly optimizes hardware systems and reconstruction algorithms simultaneously for achieving considerably better performance. However, the limited parameter space of traditional modulation elements limits the performance of the end-to-end system optimization framework. Metasurface elements are expected to bring new developments to the field of computational imaging in combination with the end-to-end system optimization framework, thanks to their wide range of micronano-structure designing parameter space and flexible multidimensional light-field modulation capabilities.

3 Metasurface-Based Computational Imaging

The exceptional capability and flexibility of metasurface in modulating multidimensional light fields make them suitable for a wide range of novel imaging applications, which further facilitates the observation of abundant light–matter interaction characteristics. In this section, we examine existing computational-enhanced metasurface imagers, categorizing them in terms of different light-field dimensions.

3.1 Spectrum

In this section, we delve into computational imaging systems that leverage the spectral modulation capabilities of metasurfaces. Specifically, we discuss the existing work in two manifolds, i.e., achromatic imaging and spectral imaging.

3.1.1 Achromatic imaging

Achromatic imaging aims to focus light of various wavelengths onto a single focal plane, employing two primary methodologies: canonical-designed phase mask methods and the inversely designed phase mask approach. Within the scope of the canonical phase mask methods, heuristic phase patterns, such as the cubic phase plate,[26] logarithmic-sphere phase,[27] shifted-axicon phase,[28] and the squared cubic phase,[29] are utilized. As shown in Fig. 2(a), Colburn et al.[30] proposed metasurface-based cubic

phase mask implementation to realize achromatic visible imaging. Since asymmetric artifacts are commonly introduced in cubic phase masks, axially symmetric masks, such as logarithmic-aspherical,[27] shifted axicon masks,[28] and square cubic phase masks,[29] are introduced to mitigate the artifacts. Figure 2(b) illustrates the imaging performance comparison of different metasurface phase masks for achromatic imaging.[31] These phase designs play a pivotal role in expanding the depth of field (DoF) across different colors and wavelength channels, creating an overlapping zone that achieves achromatism. Specifically, the logarithm sphere phase[27] extends the DoF by introducing a logarithmic phase distribution in optical systems, which divides the phase mask into annular zones with continuously varying focal lengths and allows light to maintain focus over a wider range. In addition, the shifted axicon phase[28,31] enables the creation of Bessel beams with long, nondiffracting focal lines, increasing the DoF by maintaining focus over extended distances. Furthermore, the squared cubic phase[29] is developed to effectively reduce asymmetric distortions, ensuring higher image quality, and allow extended depth of field (EDoF) imaging over longer distances. Through these phase patterns for point spread function (PSF) engineering, images exhibit approximately uniform blurring across the EDoF over different wavelengths, allowing extraction of an achromatic image using Wiener filtering[36] or total-variation regularized optimization algorithms.[37,38]

Beyond the conventional canonical phase patterns, inverse design-based methods[32–35,39] have emerged. As shown in Fig. 2(c), a multizone dispersion engineering metasurface has been proposed[32] by maximizing the minimum value of the focusing intensity at different designed frequencies in the visible wavelength range (i.e., 470 to 670 nm). This optimization employs gradient-based local methods[40] to determine zone transition locations and phase discontinuities in the multizone metasurface. To address the high computational cost in designing large-scale aperiodic metaoptics, a general computational framework for efficient optimization-based inverse design of metasurface has been proposed,[39] enabling a few minutes of computation for 2D inverse problems. To extend inverse design methods to three-dimensional (3D) Maxwell's equation, an efficient inverse design framework for aperiodic, large-scale, and 3D metasurface with fabrication constraint has been proposed.[33] As shown in Fig. 2(d), leveraging the inverse design method, both polarization-insensitive RGB-achromatic metalenses and polychromatic metalenses have been achieved, featuring aperture diameters of 1 and 2 mm, and numerical apertures (NAs) of 0.3 and 0.7. Moreover, based on the achromatic metasurface design, a virtual reality platform has been constructed, demonstrating the elegant performance of the RGB-achromatic meta-eyepiece. Based upon the 3D inverse design framework,[33] Bayati et al.[34] proposed realizing achromatic imaging with the objective function designed for extending the DoF of different wavelengths. Similarly, Wiener filtering or TV-regularized optimization-based deconvolution has been introduced to retrieve the achromatic image, as shown in Fig. 2(e). In addition, Wang et al.[35] built a library of pairwise data that maps the spectral phase response to the corresponding parameters of meta-atoms and proposed a multilayered perception (MLP) network to efficiently learn the forward simulator from metasurface structure parameters to the target phase profile. Based upon a heuristic method, i.e., the particle swarm optimization method,[41] the structure parameters of metasurface can be retrieved, and visible achromatic imaging can be realized, as shown in Fig. 2(f).
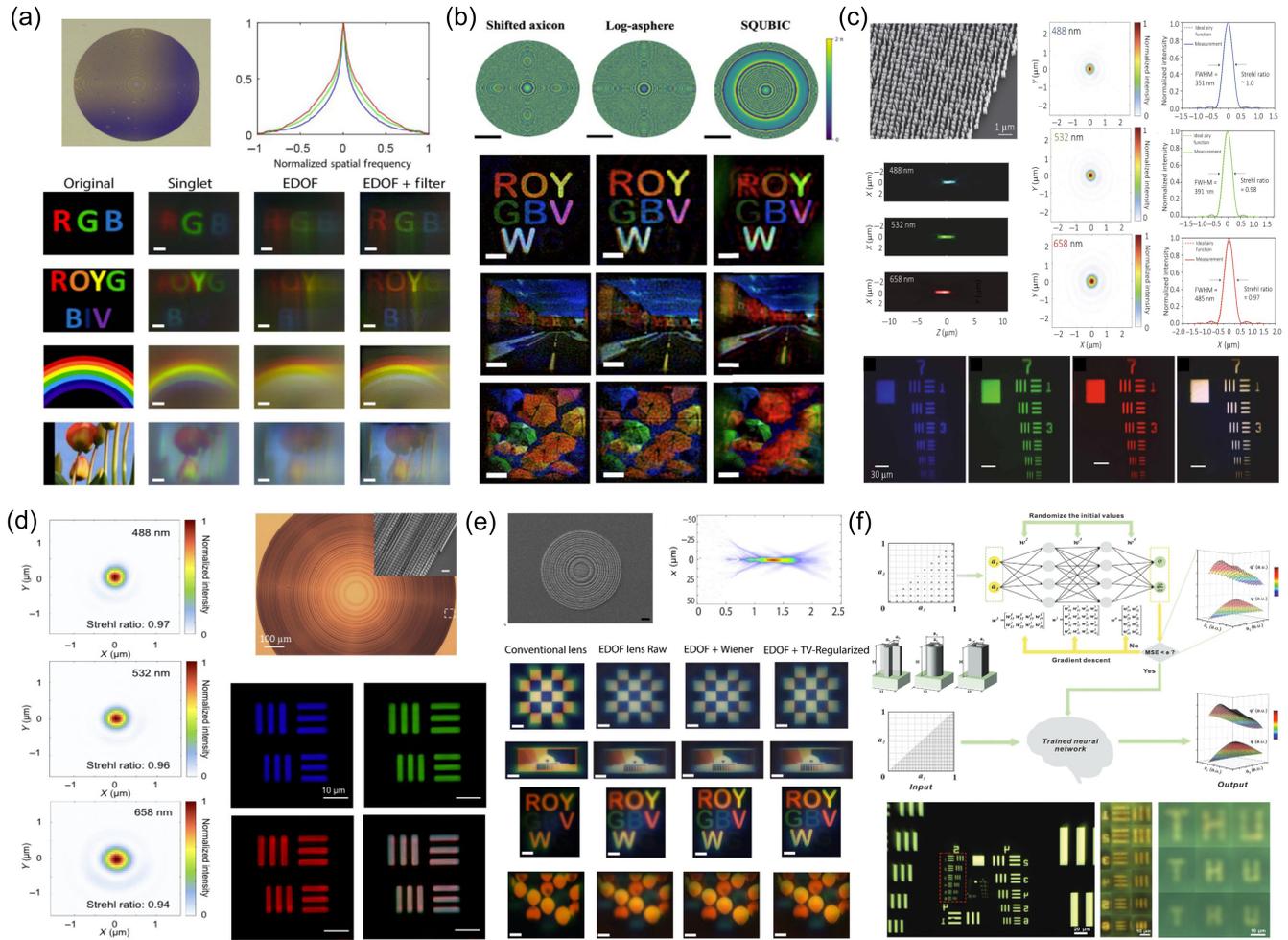
Fig. 2 Achromatic computational imaging based on metasurface modulation techniques. EDoF-based achromatic imaging: (a) cubic phase mask, reproduced with permission from Ref. 30 (CC-BY), (b) symmetric EDoF-based canonical phase mask, reproduced with permission from Ref. 31 © 2020 Chinese Laser Press. Optimization-based inverse design methods: (c) multizone dispersion-engineered metalens-based achromatic RGB focusing, reproduced with permission from Ref. 32 (CC-BY), (d) achromatic RGB imaging with efficient 3D inverse design methods, reproduced with permission from Ref. 33 (CC-BY), (e) achromatic visible imaging with EDoF-based inverse design, reproduced with permission from Ref. 31 © 2021 American Chemical Society, and (f) achromatic inverse design based upon artificial neural network, reproduced with permission from Ref. 35 © 2021 Wiley-VCH.

3.1.2 Hyperspectral imaging

Snapshot hyperspectral imaging holds significant promise across various applications, including security, remote sensing, and astronomy. A primary challenge that it faces is the substantial loss of spectral information during acquisition. Therefore, in the design of hyperspectral encoding methods, maximum spectral retention is crucial. Metasurfaces, with their versatile modulation capabilities, are emerging as powerful tools for spectral encoding in hyperspectral imaging. They offer flexibility in customizing spectral encoding patterns, making them particularly well suited for this purpose.

Among encoding schemes for snapshot hyperspectral imaging, random spectral encoding, as the simplest method, has showcased its potential to deliver high-quality and compact hyperspectral imaging.[42,43] Beyond its coding efficiency, its

resilience to manufacturing and systematic errors makes it suitable for real-world applications. Specifically, a proposal has been made for regular-shaped metasurface units to achieve random spectral encoding at the visible wavelength,[2] as shown in Fig. 3(a). By utilizing these units with randomized parameters, spectral data can be compressively encoded. To enhance encoding efficiency, the mutual coherence metric,[47] commonly used in compressive sensing theory to measure encoding efficiency, has been applied to identify improved metasurface geometries. Conventional compressive sensing (CS) reconstruction algorithms based on sparse optimization and dictionary learning are utilized to reconstruct the hyperspectral images.[48] This hyperspectral imager is employed in the detection of brain hemodynamics, specifically the spectral absorption of deoxyhemoglobin and oxyhemoglobin in rats. Furthermore, freeform-
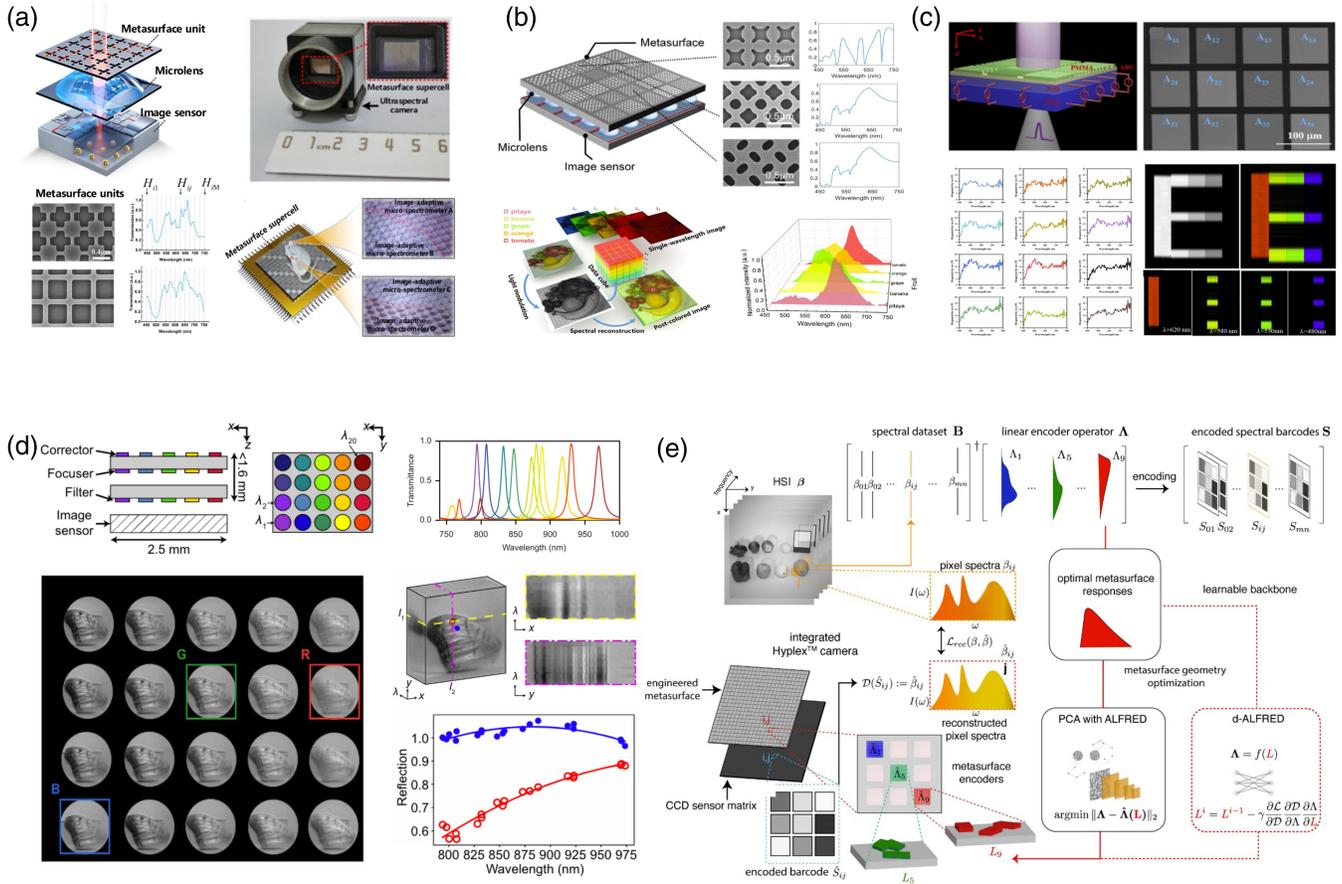
Fig. 3 Spectral modulation-based hyperspectral imaging. Hyperspectral imaging with random encoding with (a) regular-shaped metasurface design, reproduced with permission from Ref. 2 © 2022 Optical Society of America, (b) freeform-shaped metasurface, reproduced with permission from Ref. 44 © 2022 Wiley-VCH, (c) metasurface-based gratings, reproduced with permission from Ref. 45 © 2022 Optica Publishing Group, (d) multi-aperture spectral filter array-based hyperspectral imaging, reproduced with permission from Ref. 46 (CC-BY), (e) end-to-end learned optimal metasurface design, reproduced with permission from Ref. 3 © 2022 IEEE.

shaped units are introduced to realize random hyperspectral encoding with higher spatial and spectral resolution, in the visible wavelength range from 450 to 850 nm[44] [Fig. 3(b)]. Various processing techniques, from blurring to binarization, are deployed to generate these distinctive metasurface structures. The introduction of freeform units results in finer spectral domain features, ultimately preserving more spectral information. To decode the encoded image, an MLP network is utilized, offering enhanced fidelity and spectral resolution. In contrast to the aforementioned metasurface units, an all-dielectric grating-based metasurface is introduced, enabling hyperspectral encoding in the range of 400 to 800 nm,[45] as shown in Fig. 3(c). Paired with a sparse Toeplitz-basis-based spectral reconstruction algorithm, this approach achieves a spectral resolution of 4.8 nm, as confirmed by experiments.

In addition to these coding-based hyperspectral imaging methods, multiaperture spectral array-based methods have been proposed,[46] as shown in Fig. 3(d). Each channel is composed of a metasurface-based spectral filter and achromatic doublet, enabling the capturing of each spectral channel. To correct different color channels to the same viewpoint, computational parallax correction methods are implemented based on affine

transform. To further enhance encoding efficiency, learning-based spectral encoding methods are proposed. Specifically, the principal component analysis (PCA)-based hyperspectral metasurface imager, known as HyplexTM,[3] is presented, as shown in Fig. 3(e). Target transmission functions of the metasurface encoder are obtained using unsupervised learning via PCA.[49] The artificial intelligence-driven autonomous learning framework for rule-based evolutionary design (ALFRED)[50] is used for inversely designing the nanostructures of the metasurface encoder. In addition to the PCA encoding scheme, the end-to-end optimization method based on a differential extension of ALFRED (d-ALFRED)[3] is also applied to optimize the metasurface geometry and hyperspectral reconstruction simultaneously. A deep-neural-network-based forward model is used,[50] allowing for end-to-end optimization of both the encoder function of the metasurface and the subsequent reconstruction algorithm for hyperspectral segmentation. Moreover, joint optimization of the metasurface phase array and reconstruction algorithms has been explored to further enhance hyperspectral imaging.[51] A deep-unfolding neural network based upon the alternating gradient descent convex optimization framework[52,53] is implemented for image reconstruction. In addition to these

transmissive hyperspectral information encodings, a specifically designed multiwavelength off-axis focusing metamirror is proposed to capture four images reflectively in a snapshot. Through combining convex optimization and deep learning, 18 channels of hyperspectral image are retrieved from the captured image with a small amount of data for training the neural network.[54]

In summary, by fully harnessing the inherent redundancy in hyperspectral images and integrating advanced computational imaging techniques, there is optimistic anticipation that high-quality hyperspectral imaging retains essential spectral nuances that can be promisingly realized.

3.2 Polarization

In the realm of computational imaging, metasurfaces have risen as highly versatile tools, primarily owing to their distinctive capability to modulate the polarization of light. This section provides a brief overview of computational imaging techniques that harness metasurfaces' polarization modulation capabilities. We will delve into three key paradigms: polarization multiplexing, polarization routing, and polarization filtering.

3.2.1 Polarization multiplexing

Metasurfaces, with the exceptional ability to enable polarization multiplexing, provide an unprecedented degree of independent modulation freedom that greatly enhances various innovative computational imaging tasks. These tasks include the introduction of an additional encryption key,[55] underwater descattering for depth and intensity imaging,[56] on-chip wide FoV microscopic imaging,[57,58] full-Stokes imaging,[4,5] extreme DoF imaging,[59] and monocular 4D imaging.[60]

For instance, polarization multiplexing of left and right circular polarization states using metasurfaces can introduce additional layers of security keys in ghost-imaging-based information encryption applications,[55] as shown in Fig. 4(a). In addition, this meticulous tailoring of the metasurface responses to distinct light chirality enables the realization of an extreme DoF, spanning from just 3 cm to a staggering 1.7 km,[59] as shown in Fig. 4(b). A lightweight multiscale convolutional neural network (CNN) is proposed to efficiently eliminate various aberrations in the metalens array, achieving diffraction-limited resolution over nearly the entire DoF range. No color information is sacrificed for extending the DoF, and full-color extreme DoF light-field imaging could be achieved simultaneously.

Moreover, polarization multiplexing with metasurface facilitates wide FoV microscopy imaging. Typically, wide FoV microscopy relies on a microlens array with an ultracompact size, but the blind area between the adjacent microlenses can significantly hinder its practicability. By utilizing polarization multiplex with metalens array, blind areas under different incident light conditions with varying light chirality,[57,58] as shown in Figs. 4(c) and 4(d), such as left circular polarized light and right circular polarized light, can complement each other. This enables the elimination of blind areas through computational image stitching. Additionally, by multiplexing linearly orthogonal polarization states (0 deg and 90 deg) of light, depth information and high-contrast images can be recovered with reduced underwater scattering effects,[56] as shown in Fig. 4(e). Furthermore, by polarization multiplexing each sidelobe of the double-helix PSF (DH-PSF) with linearly orthogonal polarization states, 4D information, including the space, polarization, and depth

information, can be encoded and reconstructed using a physically interpretable image retrieval algorithm,[60] as shown in Fig. 4(f).

3.2.2 Polarization routing

In the realm of optical polarization imaging, the concept of polarization routing has emerged as a compelling approach, alongside polarization multiplexing. This paradigm seeks to model the polarization-dependent light-modulation capabilities of metasurfaces for the realization of full-Stokes polarization imaging. To this end, the framework of matrix Fourier optics has been introduced, representing an advancement beyond the conventional scalar Fourier transform in Fourier optics,[4] as shown in Fig. 4(g). The incorporation of the matrix Fourier transform is pivotal, as it enables the comprehensive modeling of metasurface-based polarization dimension modulation. Building upon the matrix Fourier optics theory, an inverse-design strategy of the metasurface is presented to realize the full-Stokes polarization imaging. To optimize the specifications of metasurface for effective polarization routing, the gradient descent algorithm with Lagrange multipliers methods[62] is employed. The four vertices of a tetrahedron within the Poincaré sphere are employed as the target analyzer states for polarization routing. Subsequently, the ultimate image, containing full-Stokes polarization channels, is reconstructed on a pixel-by-pixel basis via straightforward matrix inversion techniques. Notably, it is imperative to acknowledge that the utility of matrix Fourier optics modeling transcends its application in full-Stokes imaging, offering a versatile solution for the design of metasurfaces catering to an array of polarization-dependent applications. In addition to this method, an array of micropolarization routing metalenses is introduced as a solution to achieve high-sensitivity polarization imaging,[61] as shown in Fig. 4(h). Compared with the traditional filtering-based polarization imaging methods, the polarization routing schemes could achieve much higher light throughput, thus facilitating polarization imaging with high sensitivity.

3.2.3 Polarization filtering

In light of the enhanced light throughput inherent to the polarization routing-based method, it is important to acknowledge the potential limitation posed by the persistent presence of zero-order diffraction components, which can significantly impede the spatial resolution of resulting images. In response to this challenge, filtering-based polarization imaging systems use polarizing filters, such as linear polarizers or wave plates,[5,63] to directly select the desired polarization state of light for imaging. Since filtering-based polarization imaging systems operate by absorption and transmission without altering the path of light via diffraction, there are no zero-order (or any order) diffraction components involved. Specifically, high polarization extinction ratios of linear and circular polarization filters based on double-layer metallic gratings and hybrid chiral metasurfaces are proposed and realized on-chip full-Stokes imaging with ultracompactness.[63] Due to the inherent spatial-division working principle, the spatial resolution is directly compromised for the four linear and two circular polarization channels. In addition, the design principle of utilizing high extinction ratios of polarization filters requires high alignment accuracy. In addition to these polarization filter-based methods with high extinction ratios, Fan et al.[5] introduced an approach, specifically a random polarization encoding-based method, for full-Stokes imaging.
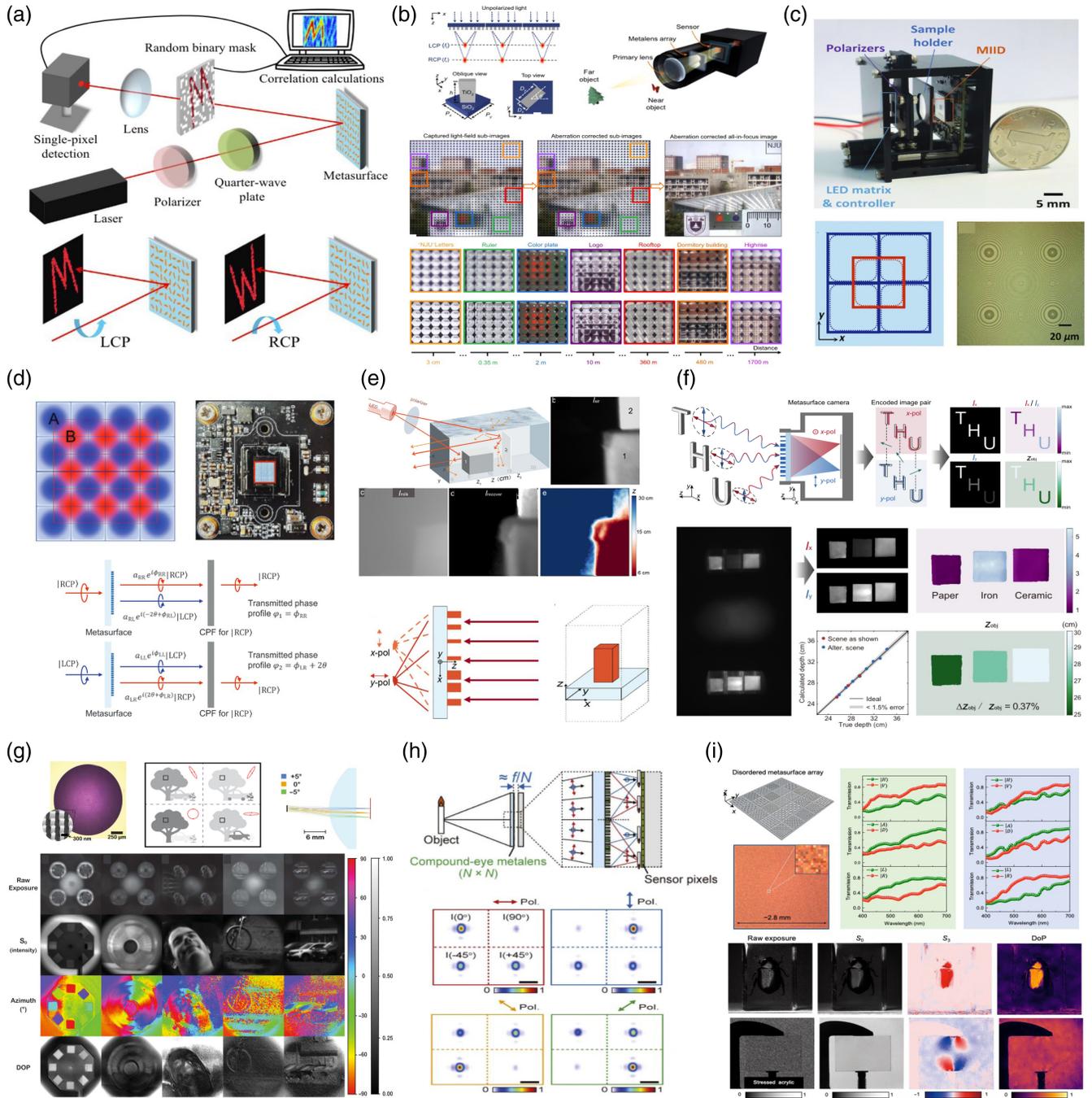
Fig. 4 Polarization modulation-based computational metasurface imager. (a) Polarization multiplexing-based single-pixel imaging, reproduced with permission from Ref. 55 (CC-BY), (b) extreme-DoF imaging with polarization multiplexing, reproduced with permission from Ref. 59 (CC-BY), (c), (d) wide FoV microscopic imaging methods with polarization multiplexing, reproduced with permission from Ref. 58 (CC-BY) and Ref. 57 (CC-BY), (e) polarization multiplexing for underwater descattering, reproduced with permission from Ref. 56 © 2021 Wiley-VCH, (f) polarization multiplexing-based 4D imaging, reproduced with permission from Ref. 60 (CC-BY), (g) full-Stokes imaging, reproduced with permission from Ref. 4 © 2019 AAAS, (h) efficient polarization imaging with polarization splitting, reproduced with permission from Ref. 61 © 2020 Optical Society of America, (i) compressive polarization imaging with random weak dichroism metasurface, reproduced with permission from Ref. 5 (CC-BY).

This method holds promise for achieving compressive sensing of full-Stokes imaging, particularly in scenarios characterized by weak dichroism, as shown in Fig. 4(i). A disordered metasurface array is proposed to serve as an efficient compressive sampling matrix in both spatial and polarization dimensions. Specifically, the transmissions for different polarization states and pixels are designed to be randomly distributed, serving as the compressive encoding for high-dimensional full-Stokes polarization images. Through calibrating the compressive sampling matrix of the metasurface, compressive sensing-based techniques can be utilized for recovering full-Stokes polarization images from the compressed measurements. Furthermore, akin to the advantages offered by random hyperspectral coding methods, the inherent robustness against manufacturing and systematic errors endows random polarization encoding schemes with practical applicability. Leveraging the attributes of weak dichroism, such schemes can surpass the light throughput of traditional polarization filtering methods. Notably, with a properly calibrated polarization response, a mask-aware full-Stokes reconstruction neural network is proposed to elegantly reconstruct the complete full-Stokes polarization image. Contrary to conventional expectations, the introduction of compressive sensing proves instrumental in alleviating design constraints associated with metasurfaces, thereby demonstrating effectiveness in realizing compact implementations of full-Stokes polarization imaging techniques.

3.3 Depth/Angle

Owing to the advantages of compactness and the subwavelength modulation capabilities, it is feasible to enable PSF engineering for depth imaging and wide-angle modulation through meticulous engineering of the phase distribution in metasurfaces. The ability to manipulate light at subwavelength scales allows metasurfaces to control light with high precision, which is critical for both depth and wide-angle imaging applications. For depth imaging, metasurfaces enable a compact, snapshot approach by ingeniously engineering the PSF. This engineered PSF is a key to creating distinct depth cues: it introduces a spatially varying blurring effect in the captured image, where the degree of blur varies with the depth of different objects. Such variations are then decoded using depth-retrieval algorithms, enabling the extraction of precise depth information from a single image. In the context of wide-angle imaging, metasurfaces offer substantial benefits due to their ability to manipulate light over a large range of angles. This is made possible by their unique property of modulating light at a scale smaller than its wavelength and provides a flexible and compact platform for capturing or projecting light across wide angles. Such capabilities are of utmost importance in the context of depth and 3D imaging. In this section, we undertake a comprehensive review of the current state of metasurface-based depth and wide-angle imaging, focusing on the incorporation of PSF engineering and wide-angle modulation techniques.

3.3.1 PSF engineering

Due to their inherent compactness, metasurfaces have emerged as crucial components for modulating PSFs in response to depth variations, thereby facilitating the encoding of depth and 3D information. This innovation holds the potential to revolutionize compact depth imaging techniques based on PSF engineering. In the scene, the depth of different points varies, and consequently, the wavefront of light emanating from these points and reaching the aperture plane differs, directly correlating with depth and represented as $U(x, y, z)$. Thus even though the phase distribution of the metasurface on the aperture plane $\phi(x, y)$ is fixed, the PSF corresponding to points at different depths captured on the sensor can be expressed as $\text{PSF}(x, y, z) = |\mathcal{F}\{U(x, y, z)\phi(x, y)\}|^2$, where $\mathcal{F}\{\cdot\}$ denotes Fourier transform and $|\cdot|^2$ denotes the complex modulus operation. Therefore, by designing the phase distribution on the aperture plane to ensure that the PSFs of different depths are as distinct as possible, depth information can be inferred by analyzing the different PSFs. One commonly employed technique for passive 3D sensing is the DH-PSF, characterized by two focal points that revolve around the center axis, with their rotation angles varying according to depth.[64] The double-helix phase profile is generated as the strategic selection of Gauss–Laguerre (GL) modes from a line in the GL modal plane. The linear relationship facilitates the creation of a clear and structured interference pattern, achieving significant DoF extension. The PSF's rotation angle exhibits a direct relationship with the depth, which allows us to deduce the depth of individual point sources based on the PSF pattern observed. Traditionally, the DH-PSF is achieved using DOE or SLM. As shown in Fig. 5(a), Jin et al.[65] introduced an all-dielectric metasurface-based DH-PSF design tailored for single-shot depth imaging, particularly in the 1500 nm wavelength range. This design exhibits significant potential for compact 3D imaging. Depth retrieval is accomplished by raster scanning a local field image using a small (128 pixels × 128 pixels) image patch, and subsequently analyzing the image spectrum to derive the local orientation of the DH-PSF. A precalibrated correlation between depth and rotation angle enables the extraction of depth information based on the inferred orientation angle. Empirical demonstrations underscore the potential of this all-dielectric DH-PSF design in imaging objects with depths ranging between 540 and 730 mm, highlighting its capacity to drive the next generation of ultracompact 3D imaging systems.

Beyond DH-PSF, the cubic phase metasurface is introduced,[26] either as a static configuration[66] [Fig. 5(b)] or reconfigurable through electrical modulation,[69] to facilitate the extraction of depth information and in-focus images with EDoF. The clean image can be recovered from the image of the cubic phase plate, using regularized filter-based deconvolution algorithms,[38] and the DH-PSF distribution can be retrieved by deconvolving the image of the DH-PSF aperture with the restored clean image through Wiener filtering. Depth information can then be retrieved from the rotation angle of the DH-PSF distribution.

As shown in Fig. 5(c), inspired by the visual mechanics of jumping spiders, which utilize defocus for depth perception, Guo et al.[58] proposed spatially multiplexing two metalenses, each with distinct defocus properties. This configuration produces two divergent defocused images in the sensor plane, allowing for depth data extraction through a lightweight depth extraction algorithm. In addition to singular-aperture-based depth-encoding techniques, there is a growing interest in metasurface array methodologies, such as the use of three closely packed hexagonal metalenses for precise 3D positioning,[68] as shown in Fig. 5(d). Incorporating a cross-correlation-based gradient descent algorithm, image plane aberrations can be corrected, enhancing the accuracy of 3D positioning.

In summary, these methodologies leverage phase encodings through PSF engineering, particularly through PSF shape variations, to extract depth data from captured images.[66,67,69] By
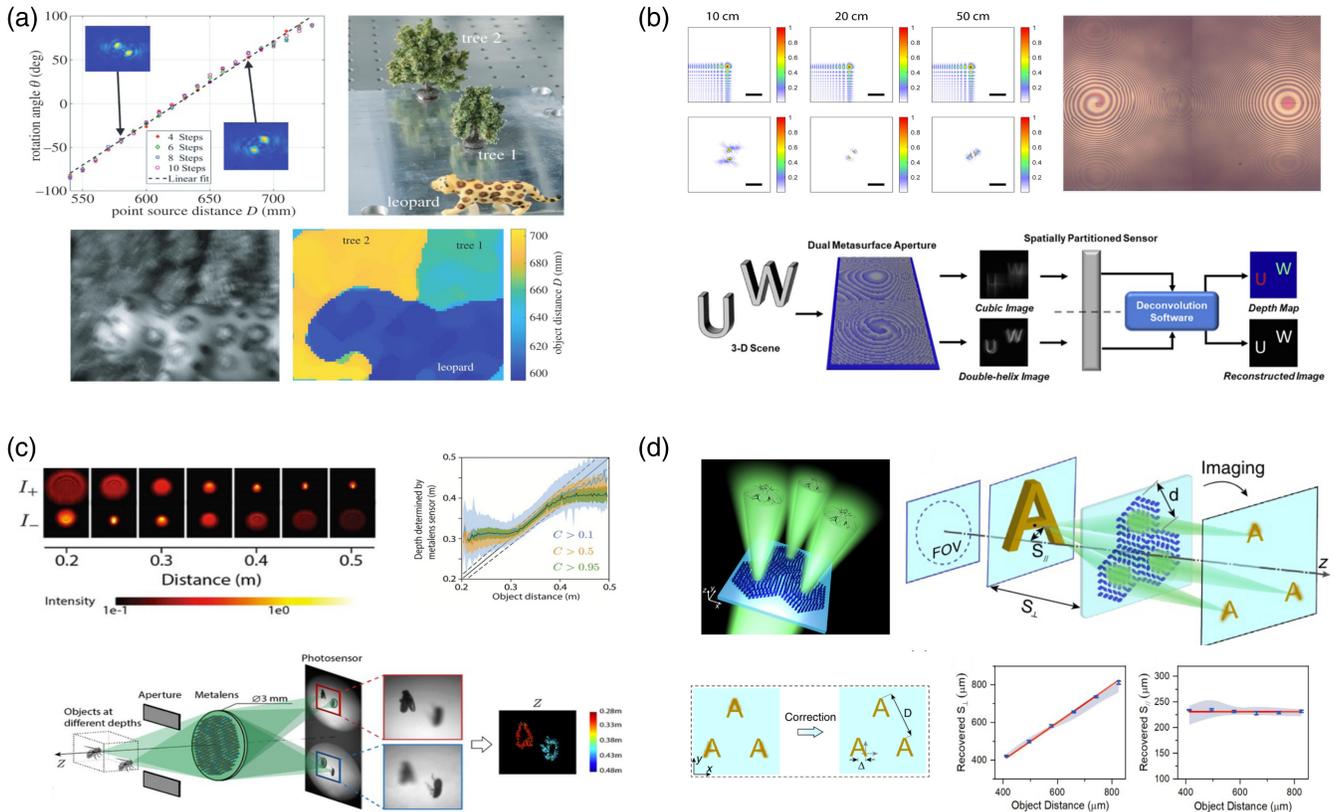
Fig. 5 Depth modulation and imaging techniques with metasurface imager. (a) DH-PSF engineering-based depth imaging, reproduced with permission from Ref. 65 (CC-BY), (b) EDoF and DH-based PSF engineering, with side-by-side metalens, reproduced with permission from Ref. 66 © 2020 ACS, (c) depth from dual-defocus multiplexing, inspired by jumping spider vision, reproduced with permission from Ref. 67 (CC-BY), and (d) triple metalens based 3D positioning, reproduced with permission from Ref. 68 © 2020 Optica Publishing Group.

calibrating the relationship between depth and PSF, depth can be extracted once the PSF of each image region is estimated using computational algorithms.

3.3.2 Wide-angle modulation

The pursuit of miniature wide-angle imaging systems, akin to compound eyes, has resulted in notable advances in recent years. An innovative approach involves an angle-sensitive ensemble of metallic plasmonic nanostructures coated onto a standard image sensor array, enabling the demonstration of 150 deg wide-angle imaging,[70] as shown in Fig. 6(a). This development obviates the need for intricate optical lenses, such as fish-eye lenses or extensive microlens arrays, as well as bulky system architectures. For image reconstruction, the truncated singular value decomposition technique is deployed to computationally fuse angle-wise sampled information to the target image.[76]

Beyond pixel-level metasurface-based angle-sensitive designs, a linear array of silicon nitride metalenses, among which each metalens possesses different phase responses to different incident angles, is integrated onto the front of a complementary metal-oxide-semiconductor image sensor, thereby enabling wide-angle imaging,[71] as shown in Fig. 6(b). Leveraging the versatility of metalens design, phase-shift configurations are implemented to ensure optimal imaging quality across specific angle distributions. Through computational stitching techniques,

images captured by different metalenses are filtered using different masks indicating the best focusing angle range and then are seamlessly combined. This approach results in the achievement of more than a 120 deg horizontal extensive viewing angle in a compact planar camera setup. To address challenges in fabricating large-aperture metalenses, researchers are turning to synthetic aperture methods (SAMs). The fundamental principle of SAM for large-aperture imaging is based on the concept of combining multiple observations taken from different viewpoints to simulate the imaging capability of a much larger aperture than what is physically possible with a single observation. This technique, widely used in radar and sonar imaging systems, allows for high-resolution imaging despite having a physically smaller sensor or aperture. As shown in Fig. 6(c), SAM-based metalens-integrated near-infrared cameras are proposed.[59,72] The proposed SAM techniques allow for the preservation of high-frequency data across four subaperture imaging setups, with high resolution comparable to that of a large aperture (which is four times the area of a subaperture). This is achieved through commonly adopted synthetic aperture reconstruction algorithms, such as Wiener filtering[36] and the Richardson–Lucy deconvolution algorithm.[77,78] In addition, polarization-insensitive, orthogonal linearly polarized, and orthogonal circularly polarized SAMs are proposed and demonstrated with simulation.[79] To improve the MTF response at high spatial frequencies without significantly
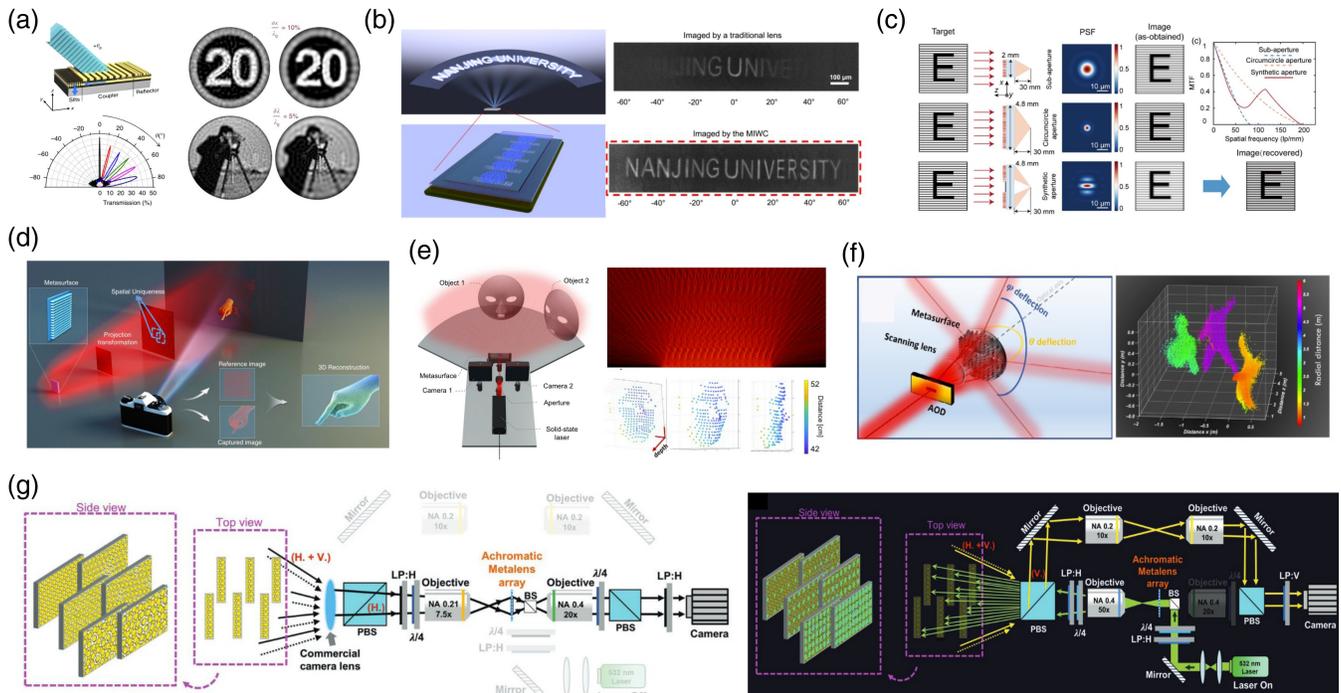
Fig. 6 Angle dimension modulation for computational imaging. Wide-angle-imaging-based upon: (a) ommatidia-inspired pixel-wise angle-sensitive filtering, reproduced with permission from Ref. 70 (CC-BY), (b) angle-selective metalens array, reproduced with permission from Ref. 71 © 2022 Optica Publishing Group, (c) synthetic aperture with four small apertures, reproduced with permission from Ref. 72 © 2021 Chinese Laser Press. Wide-angle illumination for 3D depth imaging based upon: (d) pseudo-random coding, reproduced with permission from Ref. 73 (CC-BY), (e) uniform dense light patterns, reproduced with permission from Ref. 6 (CC-BY), (f) double-zone illumination, reproduced with permission from Ref. 74 (CC-BY), and (g) dual depth imaging mode with structured light-field imaging under common light conditions and structured imaging under low-light conditions, reproduced with permission from Ref. 75 © 2022 Wiley-VCH.

reducing the efficiency at low spatial frequencies, a hybrid arrangement of high-transmission small unit cells in the high NA region of the metalens and low-transmission large unit cells in the low NA region is proposed and demonstrated to improve the maximum cutoff frequency.[80] In the domains of millimeter waves and terahertz frequencies, SAMs have been proposed to achieve high spatial resolution.[81–83] These methods utilize dynamically tunable metasurfaces to produce different radiation patterns and realize equivalent synthetic aperture imaging.

Structured illumination, a widely used technology for acquiring additional scene information through coded active illuminations, has gained prominence in depth-sensing scenarios. Depth information is obtained by illuminating scenes with structured patterns and detecting pattern distortion in captured scene images. Traditionally, structured illumination employs DOE, but limitations arise from relatively large pixel sizes, as well as constraints on angle-of-view and diffractive efficiency. The utilization of nanoscale modulation units in metasurface enables wide-angle structured illumination in a compact manner.[84] This innovation has led to submillimeter (0.24 mm) depth accuracy imaging over a depth range of 300 mm,[73] as shown in Fig. 6(d). To create unique local patterns in the entire projected illumination pattern, M-array[85] is employed as pseudo-random coding, with the Gerchberg–Saxton (GS) algorithm[86] used to calculate the required phase map of the metasurface. After capturing

images with structured illumination, a 3D reconstruction algorithm is applied, relying on the triangulation principle.[87] Similarly, metasurface-enhanced structured illumination is introduced in stereo-camera depth-sensing systems,[6] offering dense illumination with around 10^4 dots (or around 100 parallel light lines) and 180 deg angle coverage of illumination. This development has enabled depth imaging in the range of 1 m and 120 deg FoV, as shown in Fig. 6(e). The metasurface geometry for structured illumination is retrieved using the GS algorithm, and depth information is extracted through the stereo-matching algorithm, i.e., the coherent point drift algorithm.[88]

In an effort to mimic human vision, the multibeam addressing capability of metasurfaces is harnessed, and two detection schemes are proposed to enable multizone 3D imaging, simultaneously achieving scanning resolutions in the center zone of 2 deg × 2 deg and the peripheral zone of 150 deg ×150 deg, operating at the speed of 3.4 kHz (with a pixel resolution of 70 × 70), for depth imaging,[74] as shown in Fig. 6(f). A data analysis approach, along with a learning-based image retrieving algorithm, is employed to extract depth information from the captured raw data, suggesting potential for adaptive and flexible imaging. With a sophisticated design of light path, an achromatic metalens array was proposed to work under two different types of mode, depending on the illuminance intensity,[75] as shown in

Fig. 6(g). In low-light conditions, the light path is switched to the active depth-imaging mode, while under normal lighting conditions, it operates in the light-field imaging mode. A CNN facilitates the extraction of depth information in both scenarios.

In conclusion, the fusion of computational imaging techniques with the unique angle modulation capabilities of metasurfaces has paved the way for groundbreaking wide-angle illumination and imaging scenarios. These innovations hold the potential to simplify optical system designs and expand novel imaging capabilities through metasurface integration.

3.4 Compound Modulation

Leveraging the versatile multidimensional modulation capabilities of metasurfaces, complex tasks requiring cross-dimensional modulation can be significantly facilitated through metasurface-based computational imaging. Compound modulation is defined as the coordinated manipulation of various properties or dimensions of a light field to achieve specific imaging goals. This includes techniques, such as angle-spectrum modulation, where both the direction and spectrum of light are controlled, and angle-polarization modulation, which involves the simultaneous adjustment of light's angle and its polarization state, etc. These methods represent a sophisticated approach to light manipulation, allowing for more complex and refined imaging outcomes. For instance, the metasurfaces' flexible focusing depth and strong spectral dispersion capabilities allow

for the concatenation of focusing ranges of different color channels.[89] Using the proposed deep U-Net-based CNN,[90] in-focus RGB images and depth information could be retrieved, extending the DoF range, as shown in Fig. 7(a). With the flexible angle modulation capabilities of metasurfaces, ultra-compact light-field imaging is enabled with the metalens array.[96] In addition, further utilizing the strong dispersion of metasurfaces, Hua et al.[91] introduced a metalens-array with robust spectral dispersion. Employing a convex optimization-based reconstruction method, which utilizes total-variation priors in both spatial and spectral dimensions,[97] hyperspectral and light-field information could be simultaneously retrieved, as shown in Fig. 7(b).

Similarly, metasurfaces enabled the concept of color routing for color imaging.[92,93,98,99] In contrast to traditional color-filter arrays (methods like the Bayer filter), color routing offers significantly higher transmission efficiency. Heuristically designed algorithms, such as genetic algorithms,[100] are employed for inverse design of color routing and splitting, achieving optical efficiencies up to 58%, 59%, and 49% for red, green, and blue light, with an average rate of 84% across the entire visible spectrum (400 to 700 nm), roughly double the efficiency of a commercial Bayer color filter, as shown in Fig. 7(c). Ongoing efforts to further enhance this efficiency, including a path information-guided inverse design method, have pushed these numbers even higher, reaching peak efficiencies of 58.3%, 52.6%, and 69.6% for red, green, and blue light bands, as shown in Fig. 7(d).
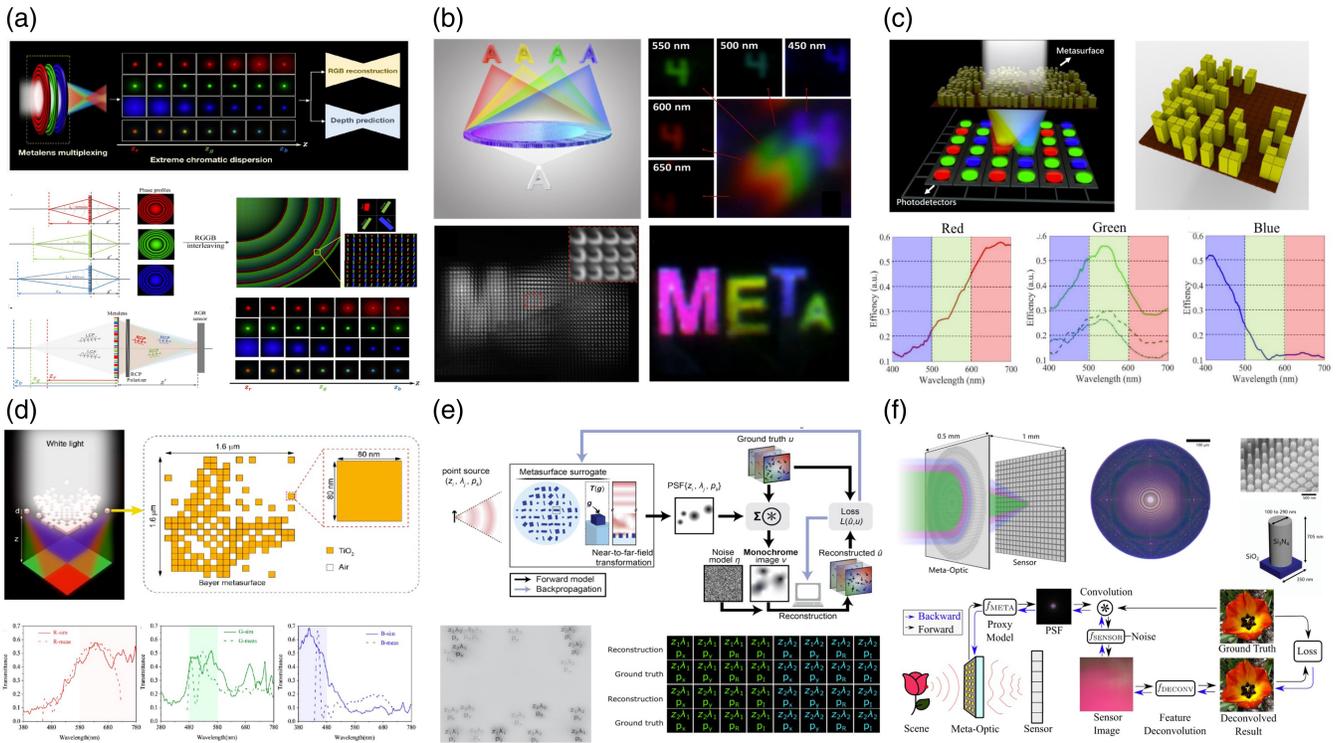

Fig. 7 Spectrum-angle/depth modulation-based computational imaging: spectral-angle joint modulation for (a) RGB and depth imaging with EDoF, reproduced with permission from Ref. 89 © 2021 ACS, (b) spectral light-field imaging, reproduced with permission from Ref. 91 (CC-BY), (c), (d) high-efficiency color imaging based upon color routing, reproduced with permission from Ref. 93 (CC-BY) and Ref. 92 © 2021 ACS, (e) single-image multichannel imaging, reproduced with permission from Ref. 94 © 2022 Optica Publishing Group, and (f) full-color wide-FoV imaging, reproduced with permission from Ref. 95 (CC-BY).

With the multidimensional modulation probability of metasurfaces, optimization-based end-to-end learning of the geometry of meta-atoms is proposed to realize single-shot multichannel imaging, including multispectral, polarization, and depth imaging,[94] as shown in Fig. 7(e). The Chebyshev interpolated surrogate model under a locally periodic approximation[39] is utilized to efficiently simulate the transmitted electric field through a large-area metasurface. An iterative conjugate-gradient method is adopted to reconstruct the target image, and the gradient for optimizing metasurface is calculated with the adjoint method.[101] Note that this kind of optimization-based end-to-end learning requires NO data set; however, the multichannel imaging capability is demonstrated only on sparsely distributed scenarios, which might be generalized to denser scenes with

Table 1 Overall summarization of existing computational metasurface imager.

	Illumination	Sensing	Reconstruction
Spectrum		• Canonical EDoF phase-based achromatic design[30,31]	• Wiener filtering[36]
		• Inversely-designed achromatic encoding[34]	• TV-regularized convex optimization algorithm[97,102,103]
		• Random hyperspectral coding with regularly-shaped,[2] free-form shaped,[44] grating-based metasurface[45]	• MLP-based spectral reconstruction[3]
		• Multiaperture hyperspectral filtering[46]	• Dictionary learning-based sparse recovery[48]
		• PCA-based spectral encoding[3]	
		• End-to-end learned metasurface design and neural network based hyperspectral retrieval and segmentation[3,51]	
Polarization		• Trilobite-inspired chiral multiplexing[59]	• Multiscale CNN[59]
		• Linear polarization multiplexing[60]	• Matrix inversion[4]
		• Inversely-designed polarization splitter with tetrahedron polarization analyzer[4]	• Gradient descent optimization algorithm[23]
		• Metalens array-based polarization splitting[61]	• Deep mask-aware compressive full-Stokes reconstruction network[5]
		• Random polarization filtering[5]	
		• Mosaic polarization filtering[63]	
Depth/Wide angle	• High-density 1D or 2D dots ($\sim 10^4$) or parallel lines (~ 100), covering 180 deg FoV[6]	• Synthetic aperture sensing[72]	• Truncated SVD technique[76]
	• M-array-based pseudo-random coding for illumination[73]	• Ommatidia-inspired pixel-wise planar wide-angle selective sensing[70]	• Wiener filtering[36] and the Richardson–Lucy deconvolution algorithm[77,78]
	• Double-zone illumination, i.e., peripheral zone (150 deg FoV) and center zone (2 deg FoV)[74]	• Angle-selective metalens linear array-based wide-angle sensing[71]	• Mask-based stitching algorithm[71]
			• Depth extraction algorithm based on triangulation,[87] stereo-matching,[88] and time-of-flight[74]
PSF engineering		• DH-PSF[65,66,69]	• Cepstrum analysis[65]
		• Jumping spider-inspired dual-defocus PSF[67]	• Depth from defocus algorithm[104]
		• Three-subaperture based on three close-packed hexagonal metalenses[68]	• Depth from parallax of correlation-aberration corrected subaperture images[68]
Compound		• Dispersion-based spectral encoding and metalens array-based light-field imaging[91]	• Convex optimization of 4D hyperspectral light-field image based on spectral–spatial sparsity prior[97]
		• Inversely-design-based color routing or splitting[92,93]	• U-Net-based RGB and depth retrieval neural network[89]
		• Tri-focus PSF over RGB channels[89]	
		• End-to-end learned metasurface design and wide FoV imaging[95] and snapshot multichannel (spectrum, polarization, and depth) imaging[94]	

deep-neural-network-based end-to-end optimization. Beyond multichannel imaging, as shown in Fig. 7(f), achromatic wide FoV imaging is proposed by combining the advantages of metasurfaces in spectral and angle modulation.[95] Using an end-to-end optimization framework, an aperture of 0.5 mm with an F-number of 2 is achieved. Notably, the response function of metasurfaces is initially fit using simpler structures like nanoposts, offering a single designable parameter, the duty cycle. However, by harnessing the more diverse modulation potentials of metasurfaces with intricate nanostructures, even higher-quality achromatic wide FoV imaging can be anticipated.

In conclusion, metasurfaces, characterized by their compactness, subwavelength, and multidimensional modulation, play a pivotal role in advancing imaging techniques. The synergy between computational methods and metasurface technology is paramount in pushing the boundaries of optical science. To provide a clearer visualization and comprehensive understanding, we have summarized this section in Table. 1.

4 Computational Imaging Framework

Different from traditional imaging, computational imaging usually introduces modulators in different locations of the imaging system with complex coding schemes for modulating the light-field information during acquisition and requires customized reconstruction algorithms for matching, as shown in Fig. 8(a). Therefore, the design framework of the computational imaging system becomes significant. Here we briefly introduce three primary design frameworks for metasurface-based computational imaging systems: the conventional imaging framework, the independent optimization-based imaging framework, and the end-to-end learned metasurface-based imaging framework.

Within the conventional imaging framework, many existing metasurface-based imaging systems are empirically designed using established patterns. Given the predictable functions of these coding patterns and modulation schemes, each module of the entire system, encompassing illumination, optical components, sensors, and postprocessing algorithms, can be independently designed with predefined input or output. Owing to its reliability and straightforward applicability, this canonical design has found extensive use in numerous computational imaging systems. However, the relatively limited modulation schemes constrain the system's capacity to approach the global optimum.

Beyond the conventional imaging framework, independently designing a metasurface with a specifically designed objective could further utilize the flexible modulation capability of metasurface. Among these methods, either the phase profile or the structure parameters of metasurface are chosen as the variables to be optimized. With the optimized metasurface for imaging, various optimization-based methods are incorporated for final target image reconstruction. As shown in Fig. 8(b), the optimal metasurface design is independently optimized without considering the reconstruction process, which might not provide direct guidance for the optimal design of metasurface toward the final target.

Recently, end-to-end optimization-based metasurface designing methods have been proposed for finding the optimal structure of the metasurface and the workflow, as shown in Fig. 8(c). Different from the independent design method, the end-to-end optimization-based metasurface design takes the reconstruction quality of the imaging target as the optimization goal. The metasurface design and the reconstruction algorithm are optimized by training the entire system in an end-to-end way to achieve the optimal solution. End-to-end joint optimization is a promising developing direction for fully exploiting the multidimensional joint-modulation potential of metasurfaces with the power of computation, potentially heralding the advent of next-generation imaging technologies.

According to the above design schemes, metasurfaces have been applied in various computational imaging systems to expand the boundaries of their capabilities. In the following, we review the existing metasurface-based computational imaging methods from two aspects, i.e., the hardware systems (including illumination and sensing) and the computational algorithms.
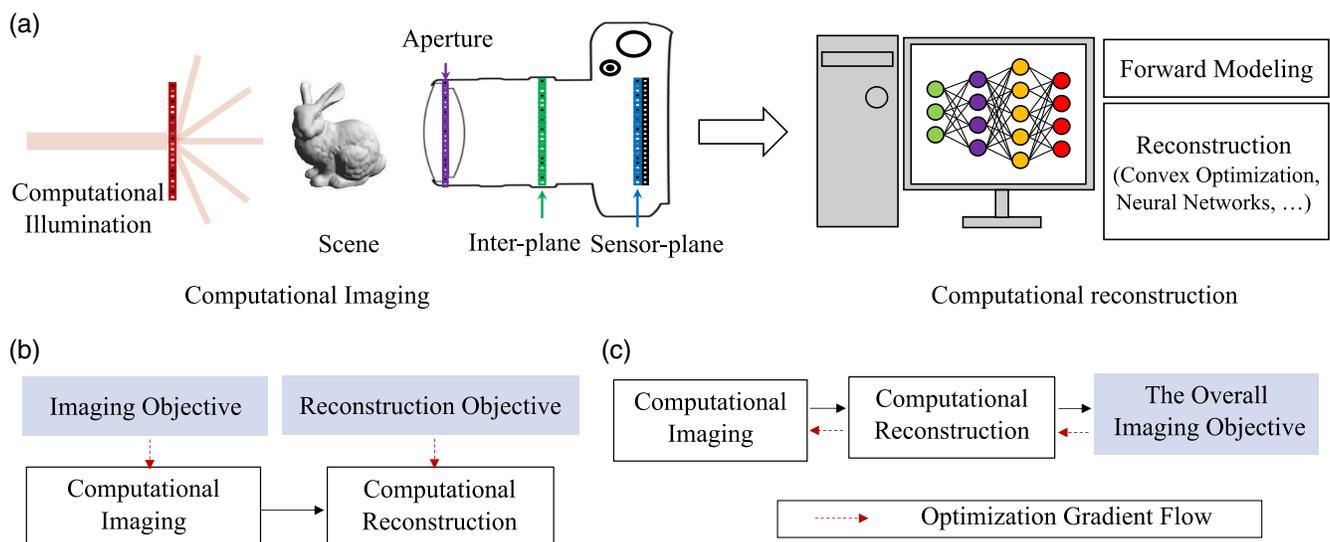

Fig. 8 Computational imaging framework of metasurface-based imaging. (a) The computational imaging process, containing optimizable imaging component and reconstruction algorithms; (b) independent optimization framework; and (c) end-to-end optimization framework.

4.1 Hardware Systems

The hardware of the computational imaging systems usually includes computational illumination and sensing modules. As for illumination, structured illumination is commonly introduced in computational imaging for a wide range of applications, including 3D imaging[105,106] and microscopic superresolution.[107] The intrinsic subwavelength modulation capabilities of metasurfaces pave the way for wide-angle modulation. This is harnessed to achieve expansive angle illumination, thereby surmounting the inefficiencies and uniformity challenges posed by diffracted beam arrays inherent to conventional optical modulators.[6,73,74] Based on the metasurface-enabled structured illumination, 3D imaging techniques could be implemented with high compactness and large angle or even full-space ranging capability. Such characteristics hold promise in a wide range of applications, such as machine vision, security protections, consumer-level 3D cameras, and virtual reality. Beyond these applications, further harnessing the unique attributes of metasurfaces to boost the existing structured illumination-based computational imaging technologies is worth exploring, providing more opportunities for the next generation of nanostructured illumination technologies.

As for the sensing module, the metasurface-based modulator could be introduced in the imaging system for implementing the designed encoding of high-dimensional light-field information. In the following, we will review the metasurface-based computational imaging works according to the plane where the modulation is introduced, i.e., the image plane modulation, the aperture plane modulation, the interspace modulation, and the lens-free modulation, as shown in Fig. 9.

Generally, image plane modulation imposes a pixel-wise optical characteristic (e.g., spectral or polarization) operation on the images, making it suitable for scenarios of color, spectral, or polarization imaging. To capture multidimensional information with a 2D sensor, various coding schemes have been proposed to encode the information into the spatial domain. Specifically, regularized periodic coding, random coding, and learning-based coding schemes are commonly employed. In the realm of conventional coding schemes, periodic coding, prevalent in many traditional imaging systems, uses a set of filters as a unit and periodically repeats the unit on the image plane, creating a mosaic-like filter array pattern. The incorporation of repeated mosaic patterns, such as the 4×4 or 5×5 spectral filter arrays found in IMEC hyperspectral sensors,[108] RGB and near-infrared spectral filter array,[109] or the 2×2 polarization filter arrays like those in Sony's Polarsens,[110] facilitates hyperspectral and polarization imaging in a snapshot manner. This approach is akin to the Bayer pattern used in RGB imaging. Demosaicking algorithms[111] could enable to restore spatial resolution across multiple hyperspectral or polarization channels. The multichannel images with spatial resolution corresponding to the original sensors could be restored. Meanwhile, the details could be lost depending on the period of the mosaic unit, the larger the period is, the more the detail loss is. In addition to periodic coding, random coding has been introduced for spatial, hyperspectral, and polarization imaging using metasurface-based random encoding. Utilizing random encoding, hyperspectral or full-Stokes polarization imaging can be achieved in a compact form, leveraging the metasurface's compactness. To enhance sampling efficiency for specific scenarios, unsupervised learning-based methods, such as PCA-based encoding, can be employed. As a widely used dimensionality reduction technique, PCA-based coding can effectively lower the dimensionality of high-dimensional data while preserving the majority of information, further boosting sampling efficiency. Beyond PCA, dimension reduction methods, such as linear discriminant analysis for supervised dimensionality reduction,[112] locally linear embedding,[113] and autoencoders[114] in deep learning, each suited to specific data types and objectives, might be further utilized for design-efficient encodings. Moreover, end-to-end learning-based encoding offers the potential for the efficient capture of high-dimensional light-field information.

Aperture plane modulation serves as another frequently adopted scheme in computational imaging. Contrasting with imaging plane modulation, aperture plane modulation achieves global modulation through convolution models. The convolutional PSF of aperture plane modulation can vary based on the wavelength, polarization, and depth of the scene points. Thus it is commonly used to capture depth, spectral, and polarization states using targeted modulation techniques like PSF engineering or to extend the DoF or FoV of imaging. Among all computational imaging techniques with metasurfaces, some metasurface design directions draw heuristic inspiration from canonical phase patterns in classical optics, such as EDoF phases and 3D localization phases. Furthermore, harnessing the advantages of metasurfaces to enhance canonical PSF-coding schemes, like multiplexing multiple canonical codings within the spectral or polarization dimensions, emerges as a promising research avenue. Additionally, when the required aperture size is impractically large, synthetic aperture-sensing methods come into play, enabling the achievement of a large metalens aperture from multiple smaller subapertures. Alongside SAMs, multiaperture-based imaging setups can also be utilized to capture higher-dimensional information, such as depth or hyperspectral data.

Beyond modulation on either the aperture or image plane, interspace modulation between these planes can effectuate local modulation across both spatial and angular domains. This method finds its niche in light-field-related imaging scenarios,
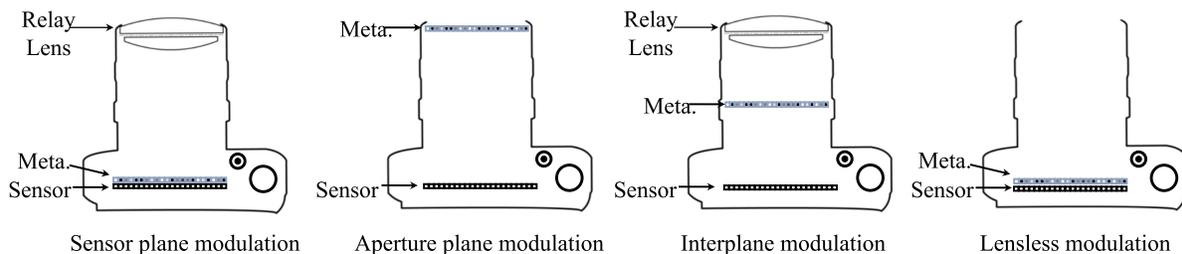

Fig. 9 Metasurface modulation-based computational sensing methods.

e.g., light-field imaging,[96] spectral light-field imaging,[91] and extreme DOF imaging.[59] In addition to the interspace modulating between the aperture and image plane, combined with the computational algorithm, the lens-free imaging methods[70] could reconstruct images from the specially designed metasurfaces without focusing lenses. Beyond angle-filtering, there exists a diverse array of lens-free techniques through introducing amplitude and phase modulations. These innovative approaches contribute to eliminating the need for traditional lenses, thereby facilitating imaging processes without the use of lenses.[10,11,115–118]

4.2 Computational Algorithms

In computational imaging systems, the target information is usually modulated in a nonexplicit manner, necessitating the specially designed algorithms to extract the information and reconstruct the desired images. The efficacy of the entire system largely hinges on the performance of these computational algorithms. Therefore, it is essential to review the existing algorithm techniques in computational imaging.

4.2.1 Overall framework

In general, the computational imaging problem consists mainly of the forward imaging process, which models the imaging process from the original signal to the measurements, and the inverse process, which reconstructs the images from measurements. As for the forward imaging process, an imaging model is commonly introduced to characterize the light-modulation function presented by the metasurface-based imaging system. This model establishes the correspondence between the essential information from the scenes and the measurements obtained. In contrast, the inverse process aims at reconstructing the target image from measurements. In this section, we will conclude the forward imaging models and reconstruction algorithms in the existing computational metasurface imaging systems.

4.2.2 Forward image modeling

Since most imaging systems follow the linear process, simple linear modeling is commonly used for simulating the physical imaging process of the computational imaging systems. Typically, matrix multiplication is commonly used to formulate the forward imaging model.[2,5,44,45] The imaging process could be represented with a measurement matrix $\mathbf{A}$, and the imaging process could be formulated as

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}, \quad (1)$$

where $\mathbf{x}$ denotes the unknown target image and $\mathbf{n}$ denotes noise. $\mathbf{y}$ denotes the measurement from the imaging system. As for the forward imaging model of spectral aberration,[30,31,34] the most simplified forward imaging model could be represented with a spatial uniform convolutional operation, i.e.,

$$\mathbf{y} = \mathbf{k} * \mathbf{x} + \mathbf{n}, \quad (2)$$

where $\mathbf{k}$ denotes the convolution kernel that represents the approximated global uniform spectral aberration, which denotes the PSF of the imaging system. The convolutional kernel could be calibrated beforehand, and clear images could be recovered with deconvolution algorithms.

In addition to the simple linear modeling, in the homogeneous, linear, isotropic matter, the solution of Maxwell equations could be simplified by the scalar Helmholtz equation $(\nabla^2 + k^2)\Psi = 0$, based on which the scalar diffraction theory could be applied for calculating the propagated light fields after passing through metasurfaces into near or far fields.[119] To characterize the effect of the modulation of metasurfaces[120,121] in the imaging process, the scalar diffraction-based modeling is widely used in the metasurface-based computational imaging systems,[122] such as the Rayleigh–Sommerfeld diffraction integral,[59] the Fresnel diffraction integral,[51] and angular spectrum methods.[73]

To account for the polarization modulation and realize the full-Stokes imaging through engineering the different diffraction order for different polarization analyzers, the matrix Fourier transform that formulates the polarization modulation effect of each unit cell of metasurface with the Jones matrix[4] is proposed, enabling the optimization of different diffraction orders of light with inverse design. Locally, the Jones matrix of the metasurface is approximated by a linearly birefringent wave plate, i.e.,

$$\mathbf{J}(\mathbf{x}, \mathbf{y}) = \mathbf{R}(\theta(\mathbf{x}, \mathbf{y})) \begin{bmatrix} e^{i\phi_x(x,y)} & 0 \\ 0 & e^{i\phi_y(x,y)} \end{bmatrix} \mathbf{R}(-\theta(\mathbf{x}, \mathbf{y})), \quad (3)$$

where $\mathbf{R}$ is the 2×2 rotation matrix, θ, φ_x, and φ_y are for the specific metasurface structures that could be easily and continuously adjusted by varying the dimensions and angular orientations of a simple dielectric pillar, which is easy to fabricate lithographically. Based upon this matrix Fourier modeling, the freedom of the polarization dimension can be optimized with inverse design methods, facilitating higher dimension modulation design of metasurfaces.

For accurately modeling the metasurface, the solution of Maxwell equations is required. However, considering the lack of a general analytical solution, the numerical solver for the Maxwell equations is commonly used for designing metasurfaces accurately. Some metasurface-based imaging or detection works[123,124] have been proposed to take the metasurface design into consideration, with different types of numerical electromagnetic solvers[125] involved in the entire computational framework, such as finite-difference time-domain (FDTD),[126] finite-difference frequency-domain (FDFD),[127] and finite-element method.[128] While with high accuracy, the required high computational cost prevents these types of forward models from wide application, especially with metasurfaces of large-scale aperture.

To enable the efficient inverse design of large-scale aperture, the semiphysical models, which overcome the computational-cost bottlenecks in pure physics-based models with fast approximate operations, have been proposed.[33,39] Specifically, a 3D fast approximate solver based on the convolution of local fields and vectorial Green function is proposed to predict the local field of an arbitrary meta-atom with fabricable parameters. Precalculating the accurate local fields of a training set of meta-atoms with rigorous coupled-wave analysis (RCWA),[39] a surrogate model based upon the Chebyshev interpolation to predict the local field of an arbitrary meta-atom is presented, achieving 6 orders of magnitude faster than direct simulation using RCWA. This type of semiphysical differential model has been utilized in efficient inverse design for achromatic RGB or polychromatic focusing with a metasurface of large aperture[33] and single-image multichannel imaging,[94] largely reducing the calculation speed of the forward imaging process.

As for the end-to-end computational imaging framework, the forward model needs to be computed repeatedly during the training process of neural networks, further exacerbating the computational cost problem of forward modeling. The empirical models, which further simplify the forward imaging process with empirical fitting models, are proposed to further speed up the model computation. Ethan et al.[95] introduced polynomial fitting to approximate the differential relationships from the resulting phase at a certain wavelength to the duty cycle and from the duty cycle of metasurface atom to the resultant phase at different wavelengths. The duty cycle and the corresponding phase mapping library for function fitting are precomputed with RCWA.[33] Based on the differential relationship between the duty cycle and the phase at different wavelengths, the PSF array of varying incident angles is calculated to simulate the blur distribution over a large FoV with different degrees of aspheric blur. This spatially varied PSF blur assumption could more accurately approximate the aberration than the global uniform blur, especially when the FoV is large, and a clean wide FoV image is recovered with the proposed neural network.

In addition to the empirical modeling methods, to realize the optimizable forward imaging process, data-driven-based methods have been proposed to train the deep neural network that maps the physical parameters of the system to the corresponding response functions. In addition, various types of neural network-based modeling methods, e.g., U-Net encoder-based feature extraction, MLP-based hyperspectral transmission projector,[3,129] and artificial neural network-based metasurface coding pattern,[130] have also been proposed to learn the differential forward propagator of metasurfaces. The differential forward model that predicts the macrography response characteristics of optical elements from the parameters of nanomicro-structures could facilitate the end-to-end optimization framework in metasurface-based computation imaging.

4.2.3 Reconstruction method

Depending on the characteristics of the algorithms, the reconstruction methods of metasurface-based computational imaging systems could be divided into the handcrafted reconstruction method, the closed-form reconstruction method, the iterative optimization reconstruction method, and the learning-based reconstruction method. In the early years of the development of computational imaging, handcrafted reconstruction methods were commonly used for their simplicity, for example, the demosaicking algorithm for filter array-based color imaging,[111] depth estimation through stereo block matching from binocular imaging,[131] and depth from defocus algorithms.[104] Although various handcrafted reconstruction methods have been applied, the lack of rigorous mathematical deviation and relatively poor performance limited the development of the handcrafted reconstruction method, and optimization-based methods were introduced.

Different from the handcrafted methods, for the optimization-based methods, explicit objective functions are formulated to be optimized. Generally, the optimization problem of computational imaging could be commonly formulated with the objective function composed of a data fidelity term and a prior term. The data fidelity term restricts the reconstructed target image to reproduce the measurement with the forward imaging model, and the prior term enforces the sparsity of the target image in a certain transform domain, such as total variation transform domain,[132] discrete cosine transform domain,[133] wavelet transform domain,[134] or statistically learned over-complete dictionary-based transform domain (dictionary priors),[135,136] etc. To solve the optimization problem, two types of algorithms, i.e., the closed-form and iterative optimization methods, can be applied. If the optimization problem has an analytic solution, i.e., the optima could be computed with a closed-form solution, the closed-form reconstruction method could be applied to calculate the target images, such as Wiener filtering.[137] In general, the closed-form reconstruction method is highly efficient, with good reconstruction quality. However, the requirement of explicit analytic solution limits the form of the objective functions, so that many problems with complex forward models or prior functions cannot be solved by closed-form methods. For handling the optimization problems without the closed-form solution, the iterative optimization reconstruction algorithms, such as the Richardson–Lucy algorithm,[77,78] FISTA,[138] TV-regularized optimization,[102,103] and dictionary learning-based sparse optimization,[48] have been introduced in the metasurface-based computational imaging methods. Notably, the optimization-based methods require NO training data sets, so that they can be applied for uncharted scenarios, e.g., the breakthrough computational microscopy or telescope imaging systems that could see scenarios that never have been imaged by mankind.

Aside from the conventional optimization-based reconstruction method, the learning-based reconstruction methods have become the research focus currently. Especially, the neural network-based reconstruction methods have attracted much interest for their remarkably improved performance in recent years compared with conventional optimization-based methods. In the reconstruction of metasurface-based imaging using neural networks, a variety of CNN architectures have been developed. Specifically, the multiscale convolutional network architecture is proposed for RGB and depth image retrieval,[89] wide FoV deconvolution,[95] and achromatic aberration removal of light-field images.[59] The multiscale CNN, notable for its efficiency in contextual and detail capture, proves particularly efficient at restoring high-fidelity image details. Its design allows for the extraction and integration of features at various scales, facilitating the accurate reconstruction of complex images. For recovering full-Stokes images from compressed measurements, the tailored deep mask-aware compressive neural network is proposed.[5] It leverages the calibrated compressive measurement matrix of the physical image system to recover full-Stokes images and reconstruct high-dimensional full-Stokes data from compressive measurements. Through training the network to be robust to the disturbance in the compressive measurement matrix, the trained neural network could be robust to noise disturbance existing in the imaging process. Beyond the CNN-based methods, the transformer-based neural network has recently emerged and shows promising performance in various image restoration tasks,[139,140] which have the potential to be applied in metasurface-based computational imaging tasks for further pushing imaging performance. Overall, these learning-based methods leverage the ability of neural networks to learn complex, nonlinear mappings from data, which is often not possible with conventional techniques. They are characterized by their adaptability to different types of data and problems, the use of large data sets for training, and their ability to generalize to new, unseen data. The success of these methods in metasurface-based imaging suggests that they can effectively handle the complex and varied distortions introduced by such optical systems, outperforming classical algorithms that might not be

as flexible or powerful. However, there still exist several challenges that need to be resolved for practical applications. Details of the challenges are discussed in Sec. 5.

5 Challenges and Perspective

The foregoing review underscores that metasurface-based computational imaging has emerged as a focal point of research and has experienced significant advancements in recent years. However, there remain intricate challenges that must be addressed to seamlessly integrate computation and metasurfaces, with the aim of achieving high-performance and practical metasurface-based computational imaging systems. In the subsequent discussion, we will delve into these challenges and suggest potential research directions to overcome these challenges.

The foremost challenge lies in constructing a differential forward imaging model that strikes a balance between accuracy and computational efficiency. Specifically, for end-to-end design, differential light-field forward propagation models need to be integrated with the reconstruction algorithm. This integration permits the gradient of the reconstruction loss to be backpropagated, thereby facilitating joint optimization of both the metasurface design and the reconstruction algorithm. However, achieving a precise calculation of the output light field postinteraction with the metasurface structures is computationally demanding. This often necessitates solving the Maxwell equation using numerical methods, such as FDTD[126] or FDFD,[127] particularly when designing a large-scale metasurface aperture that encompasses millions of metasurface units. To address this challenge, two prevailing solutions are currently employed: black-box-style learning with substantial training data[3,129] and approximation-based acceleration techniques.[33,39] In black-box learning, the internal working principles of the imaging models are not directly interpretable. Instead, the model takes input data and produces the desired output after being trained on a substantial data set. This method is powerful for complex tasks where defining an explicit imaging model is nearly impossible. Specifically, training data sets comprising metasurface structure distributions and their corresponding modulation functions are generated using high-precision numerical methods. This approach suffers from the trade-off between approximation accuracy and the computational burden of generating expansive data sets. On the other hand, approximation-based methods, which are based on a certain understanding and knowledge of the imaging process, apply approximations to some processes within the model. These methods require many fewer data compared to what is needed by black-box approaches. However, the precision of such methods depends on the errors introduced by the approximations themselves. Consequently, efficiently discerning and leveraging inherent patterns within the forward propagator to derive a reasonably accurate differential light propagator remains a pressing issue meriting thorough investigation.

The second challenge lies in translating optimized metasurface designs from simulations to tangible experimental environments. It has been observed that metasurface-based computational imaging systems might not manifest the same efficacy in real-world scenarios as they do in simulations. This disparity is attributed to the deviations in the metasurface due to fabrication errors, inexact forward modeling of the imaging process, and the complex noise interference that is intrinsic to the physical optics system. To pave a path forward, there is a pressing need to devise a comprehensive forward imaging model. Such a model would factor in physical imperfections, capture fabrication discrepancies, and more precisely, account for the noise of the physical imaging system. By tailoring the model to be resilient against these discrepancies, we can anticipate a decrease in potential performance downturns. Furthermore, once the real-world imaging system is built, recalibrating its response and meticulous refinement of the reconstruction algorithm emerges as valuable strategies to counteract performance degradation. In essence, these methodologies hold the promise of narrowing the gap between simulation forecasts and real-world experimental outcomes. Crafting a framework that ensures generalization to actual physical experiments is a research topic worthy of exploration, bearing immense practical implications.

The third challenge revolves around the limited availability of high-dimensional data sets necessary for optimizing metasurfaces within an end-to-end design framework. The multidimensional modulation potential of metasurfaces could be harnessed for efficient high-dimensional imaging. Yet, current high-dimensional data sets are decidedly scarce. For hyperspectral imaging, only a few hundred hyperspectral images have been captured across different camera setups and scenarios.[3,141–144] In the realm of polarization, the available data sets primarily feature limited linear polarization with RGB images.[145–147] The full-Stokes polarization spectral image data set has only recently been made available,[148] but it encompasses merely 63 images. To navigate this challenge, approaches like unsupervised[149] or weakly supervised learning[150] may offer solutions. Additionally, a concerted effort to capture more multidimensional light-field image data sets could alleviate this shortfall, providing significant momentum to advances in this domain.

The last challenge lies in the potential for local-optimum entrapment when optimizing the metasurface using the end-to-end design framework. At present, the end-to-end computational imaging approach predominantly supervises the terminal imaging outcome. Meanwhile, metasurface optimization takes place in the earlier layers of the neural network. This structure poses a risk: the gradient could dissipate as it propagates from the terminal loss to the metasurface layer, leading the system's optimization process into local minima.[151] Regularization and normalization techniques have been proposed to mitigate the problem.[152,153] Regularization-based methods[152] promote the optimization process to focus on more robust and generalized patterns in the data, rather than fitting to noise or specific idiosyncrasies of the training data set. In addition, normalization-based methods[153] could help in stabilizing the learning process and help to mitigate the problem of trapping into local minima. In addition, addressing the need for early layer supervision to ensure the imaging system converges toward a global optimum remains a topic of ongoing investigation. This concept involves providing additional guidance to the network to promote convergence towards practical and optimal performance.[154] Delving into this challenge through an interdisciplinary approach is crucial for the optimal development of highly efficient metasurface-based computational imaging systems.

6 Conclusion

This review offers an overview of metasurface-based imaging systems through the lens of computational imaging. Observing current research trends, we posit that leveraging the multidimensional, flexible, subwavelength, compact modulation capacities

of metasurfaces can pave the way for the development of innovative computational imaging systems, enabling the realization of practical, high-performance, and ultracompact imaging solutions. In seeking to narrow the gap between current practices and future aspirations, we identify four key challenges that merit exploration. These challenges highlight the directions for developing cutting-edge imaging techniques and harnessing the full potential of computational imaging in tandem with nanophotonics.

Code and Data Availability

Data sharing is not applicable to this article, as no new data were created or analyzed.

References

1. K. Mohammadreza et al., "Metalenses at visible wavelengths: diffraction-limited focusing and subwavelength resolution imaging," *Science* **352**(6290), 1190–1194 (2016).
2. X. Jian et al., "Dynamic brain spectrum acquired by a realtime ultraspectral imaging chip with reconfigurable metasurfaces," *Optica* **9**(5), 461–468 (2022).
3. M. Maksim et al., "Real-time hyperspectral imaging in hardware via trained metasurface encoders," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 12692–12702 (2022).
4. A. Rubin Noah et al., "Matrix Fourier optics enables a compact full-Stokes polarization camera," *Science* **365**(6448), eaax1839 (2019).
5. Q. Fan et al., "Disordered metasurface enabled single-shot full-Stokes polarization imaging leveraging weak dichroism," *Nat. Commun.* **14**(1), 7180 (2023).
6. G. Kim et al., "Metasurface-driven fullspace structured light for three-dimensional imaging," *Nat. Commun.* **13**(1), 5920 (2022).
7. M. G. L. Gustafsson, "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy," *J. Microsc.* **198**(2), 82–87 (2000).
8. M. J. Rust, M. Bates, and X. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)," *Nat. Methods* **3**(10), 793–796 (2006).
9. G. Zheng, R. Horstmeyer, and C. Yang, "Wide-field, high-resolution Fourier ptychographic microscopy," *Nat. Photonics* **7**(9), 739–745 (2013).
10. J. K. Adams et al., "In vivo lensless microscopy via a phase mask generating diffraction patterns with high-contrast contours," *Nat. Biomed. Eng.* **6**(5), 617–628 (2022).
11. B. Vivek et al., "Lensless imaging: a computational renaissance," *IEEE Signal Process Mag.* **33**(5), 23–35 (2016).
12. M. O'Toole, D. B. Lindell, and G. Wetzstein, "Confocal non-line-of-sight imaging based on the light-cone transform," *Nature* **555**(7696), 338–341 (2018).
13. A. Velten et al., "Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging," *Nat. Commun.* **3**(1), 745 (2012).
14. T. E. Bishop, S. Zanetti, and P. Favaro, "Light field superresolution," in *IEEE Int. Conf. Comput. Photogr.*, IEEE, pp. 1–9 (2009).
15. P. Llull et al., "Coded aperture compressive temporal imaging," *Opt. Express* **21**(9), 10526–10545 (2013).
16. E. Reinhard et al., *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting*, Morgan Kaufmann (2010).
17. A. Levin et al., "Image and depth from a conventional camera with a coded aperture," *ACM Trans. Graph.* **26**(3), 70-es (2007).
18. G. R. Arce et al., "Compressive coded aperture spectral imaging: an introduction," *IEEE Signal Process. Mag.* **31**(1), 105–115 (2013).
19. K. Akiyama et al., "First M87 event horizon telescope results. IV. Imaging the central supermassive black hole," *Astrophys. J. Lett.* **875**(1), L4 (2019).
20. A. Velten et al., "Femto-photography: capturing and visualizing the propagation of light," *ACM Trans. Graphics* **32**(4), 44 (2023).
21. D. J. Brady et al., "Multiscale gigapixel photography," *Nature* **486**(7403), 386–389 (2012).
22. J. N. Mait, G. W. Euliss, and R. A. Athale, "Computational imaging," *Adv. Opt. Photonics* **10**(2), 409–483 (2018).
23. S. Ruder, "An overview of gradient descent optimization algorithms," arXiv:1609.04747 (2016).
24. S. P. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press (2004).
25. G. Barbastathis, A. Ozcan, and G. Situ, "On the use of deep learning for computational imaging," *Optica* **6**(8), 921–943 (2019).
26. E. R. Dowski and W. T. Cathey, "Extended depth of field through wave-front coding," *Appl. Opt.* **34**(11), 1859–1866 (1995).
27. W. Chi and N. George, "Electronic imaging using a logarithmic asphere," *Opt. Lett.* **26**(12), 875–877 (2001).
28. Z. Zhai et al., "Extended depth of field through an axicon," *J. Mod. Opt.* **56**(11), 1304–1308 (2009).
29. N. Patwary et al., "Experimental validation of a customized phase mask designed to enable efficient computational optical sectioning microscopy through wavefront encoding," *Appl. Opt.* **56**(9), D14–D23 (2017).
30. S. Colburn, A. Zhan, and A. Majumdar, "Metasurface optics for full-color computational imaging," *Sci. Adv.* **4**(2), eaar2114 (2018).
31. L. Huang et al., "Design and analysis of extended depth of focus metalenses for achromatic computational imaging," *Photonics Res.* **8**(10), 1613–1623 (2020).
32. Z. Li et al., "Meta-optics achieves RGB-achromatic focusing for virtual reality," *Sci. Adv.* **7**(5), eabe4458 (2021).
33. Z. Li et al., "Inverse design enables large-scale high-performance meta-optics reshaping virtual reality," *Nat. Commun.* **13**(1), 2409 (2022).
34. E. Bayati et al., "Inverse designed extended depth of focus meta-optics for broadband imaging in the visible," *Nanophotonics* **11**(11), 2531–2540 (2021).
35. F. Wang et al., "Visible achromatic metalens design based on artificial neural network," *Adv. Opt. Mater.* **10**(3), 2101842 (2022).
36. R. C. Gonzales and P. Wintz, *Digital Image Processing*, Addison-Wesley Longman (1987).
37. Hunt, "Bayesian methods in nonlinear digital image restoration," *IEEE Trans. Comput.* **C-26**(3), 219–229 (1977).
38. P. Getreuer, "Total variation deconvolution using split Bregman," *Image Process. On Line*, **2**, 158–174 (2012).
39. R. Pestourie et al., "Inverse design of large-area metasurfaces," *Opt. Express* **26**(26), 33732–33747 (2018).
40. K. Svanberg, "A class of globally convergent optimization methods based on conservative convex separable approximations," *SIAM J. Optim.* **12**(2), 555–573 (2002).
41. J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. Int. Conf. Neural Networks*, IEEE, Vol. 4, pp. 1942–1948 (1995).
42. Y. Zhao et al., "Hyperspectral imaging with random printed mask," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, IEEE, pp. 10149–10157 (2019).

43. M. Yako et al., "Video-rate hyperspectral camera based on a CMOS-compatible random array of Fabry–Perot filters," *Nat. Photonics* **17**(3), 218–223 (2023).

44. J. Yang et al., "Ultraspectral imaging based on metasurfaces with freeform shaped meta-atoms," *Laser Photonics Rev.* **16**(7), 2100663 (2022).

45. Z. Wu et al., "Random color filters based on an all-dielectric metasurface for compact hyperspectral imaging," *Opt. Lett.* **47**(17), 4548–4551 (2022).

46. A. McClung et al., "Snapshot spectral imaging with parallel metasystems," *Sci. Adv.* **6**(38), eabc7646 (2020).

47. M. F. Duarte and Y. C. Eldar, "Structured compressed sensing: from theory to applications," *IEEE Trans. Signal Process.* **59**(9), 4053–4085 (2011).

48. S. Zhang et al., "A spectral reconstruction algorithm of miniature spectrometer based on sparse optimization and dictionary learning," *Sensors* **18**(2), 644 (2018).

49. M. E. Tipping and C. M. Bishop, "Probabilistic principal component analysis," *J. R. Stat. Soc. Ser. B Stat. Methodol.* **61**(3), 611–622 (1999).

50. F. Getman et al., "Broadband vectorial ultrathin optics with experimental efficiency up to 99% in the visible region via universal approximators," *Light Sci. Appl.* **10**(1), 47 (2021).

51. Q. Zhang et al., "End-to-end joint optimization of metasurface and image processing for compact snapshot hyperspectral imaging," *Opt. Commun.* **530**, 129154 (2023).

52. D. S. Jeon et al., "Compact snapshot hyperspectral imaging with diffracted rotation," *ACM Trans. Graphics* **38**(4), 117 (2019).

53. L. Wang et al., "Hyperspectral image reconstruction using a deep spatial-spectral prior," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, pp. 8032–8041 (2019).

54. C.-H. Lin et al., "Metasurface-empowered snapshot hyperspectral imaging with convex/deep (CODE) small-data learning theory," *Nat. Commun.* **14**(1), 6979 (2023).

55. H.-C. Liu et al., "Single-pixel computational ghost imaging with helicity-dependent metasurface hologram," *Sci. Adv.* **3**(9), e1701477 (2017).

56. F. Zhao et al., "Metalens-assisted system for underwater imaging," *Laser Photonics Rev.* **15**(8), 2100097 (2021).

57. X. Ye et al., "Chip-scale metalens microscope for wide-field and depth-of-field imaging," *Adv. Photonics* **4**(4), 046006 (2022).

58. B. Xu et al., "Metalens-integrated compact imaging devices for wide-field microscopy," *Adv. Photonics* **2**(6), 066004 (2020).

59. Q. Fan et al., "Trilobite-inspired neural nanophotonic light-field camera with extreme depth-of-field," *Nat. Commun.* **13**(1), 2130 (2022).

60. Z. Shen et al., "Monocular metasurface camera for passive single-shot 4D imaging," *Nat. Commun.* **14**(1), 1035 (2023).

61. M. Miyata, M. Nakajima, and T. Hashimoto, "Compound-eye metasurface optics enabling a high-sensitivity, ultra-thin polarization camera," *Opt. Express* **28**(7), 9996–10014 (2020).

62. D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*, Academic Press (2014).

63. J. Zuo et al., "Chip-integrated metasurface full-Stokes polarimetric imaging sensor," *Light Sci. Appl.* **12**(1), 191 (2023).

64. A. Greengard, Y. Y. Schechner, and R. Piestun, "Depth from diffracted rotation," *Opt. Lett.* **31**(2), 181–183 (2006).

65. C. Jin et al., "Dielectric metasurfaces for distance measurements and three-dimensional imaging," *Adv. Photonics* **1**(3), 036001 (2019).

66. S. Colburn and A. Majumdar, "Metasurface generation of paired accelerating and rotating optical beams for passive ranging and scene reconstruction," *ACS Photonics* **7**(6), 1529–1536 (2020).

67. Q. Guo et al., "Compact single-shot metalens depth sensors inspired by eyes of jumping spiders," *Proc. Natl. Acad. Sci. U. S. A.* **116**(46), 22959–22965 (2019).

68. W. Liu et al., "Aberration-corrected three-dimensional positioning with a single-shot metalens array," *Optica* **7**(12), 1706–1713 (2020).

69. Z. Ma et al., "Reconfigurable metalens with phase-change switching between beam acceleration and rotation for 3D depth imaging," *Micromachines* **13**(4), 607 (2022).

70. L. C. Kogos et al., "Plasmonic ommatidia for lensless compound-eye vision," *Nat. Commun.* **11**(1), 1637 (2020).

71. J. Chen et al., "Planar wide-angle-imaging camera enabled by metalens array," *Optica* **9**(4), 431–437 (2022).

72. F. Zhao et al., "Synthetic aperture metalens," *Photonics Res.* **9**(12), 2388–2397 (2021).

73. X. Jing et al., "Single-shot 3D imaging with point cloud projection based on metadevice," *Nat. Commun.* **13**(1), 7842 (2022).

74. R. J. Martins et al., "Metasurface-enhanced light detection and ranging technology," *Nat. Commun.* **13**(1), 5724 (2022).

75. M. K. Chen et al., "A meta-device for intelligent depth perception," *Adv. Mater.* **35**, 2107465 (2022).

76. P. C. Hansen, T. Sekii, and H. Shibahashi, "The modified truncated SVD method for regularization in general form," *SIAM J. Sci. Stat. Comput.* **13**(5), 1142–1150 (1992).

77. W. H. Richardson, "Bayesian-based iterative method of image restoration," *J. Opt. Soc. Am.* **62**(1), 55–59 (1972).

78. L. B. Lucy, "An iterative technique for the rectification of observed distributions," *Astron. J.* **79**, 745 (1974).

79. J. Hu et al., "Polarization-insensitive, orthogonal linearly polarized and orthogonal circularly polarized synthetic aperture metalenses," *Photonics* **10**(4), 348 (2023).

80. C.-Y. Yu et al., "Optical performance of synthetic aperture metalens based on hybrid unit-cells," *Mater. Sci. Semicond. Process.* **170**, 107982 (2024).

81. J. N. Gollub et al., "Large metasurface aperture for millimeter wave computational imaging at the human-scale," *Sci. Rep.* **7**(1), 42650 (2017).

82. J. Hunt et al., "Metamaterial apertures for computational imaging," *Science* **339**(6117), 310–313 (2013).

83. M. Boyarsky et al., "Synthetic aperture radar with dynamic metasurface antennas: a conceptual development," *J. Opt. Soc. Am. A* **34**(5), A22–A36 (2017).

84. Y. Ni et al., "Metasurface for structured light projection over 120 field of view," *Nano Lett.* **20**(9), 6719–6724 (2020).

85. J. F. MacWilliams and N. J. A. Sloane, "Pseudo-random sequences and arrays," *Proc. IEEE* **64**(12), 1715–1729 (1976).

86. R. W Gerchberg, "A practical algorithm for the determination of plane from image and diffraction pictures," *Optik* **35**(2), 237–246 (1972).

87. O. Faugeras, *Three-Dimensional Computer Vision: A Geometric Viewpoint*, MIT Press (1993).

88. A. Myronenko and X. Song, "Point set registration: coherent point drift," *IEEE Trans. Pattern Anal. Mach. Intell.* **32**(12), 2262–2275 (2010).

89. S. Tan et al, "3D imaging using extreme dispersion in optical metasurfaces," *ACS Photonics* **8**(5), 1421–1429 (2021).

90. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," *Lect. Notes Comput. Sci.* **9351**, 234–241 (2015).

91. X. Hua et al., "Ultra-compact snapshot spectral light-field imaging," *Nat. Commun.* **13**(1), 2732 (2022).

92. J. Li et al., "Single-layer Bayer metasurface via inverse design," *ACS Photonics* **9**(8), 2607–2613 (2022).

93. X. Zou et al., "Pixel-level Bayer-type colour router based on metasurfaces," *Nat. Commun.* **13**(1), 3288 (2022).

94. Z. Lin et al., "End-to-end metasurface inverse design for single-shot multi-channel imaging," *Opt. Express* **30**(16), 28358–28370 (2022).

95. E. Tseng et al., "Neural nano-optics for high-quality thin lens imaging," *Nat. Commun.* **12**(1), 6493 (2021).

96. R. J. Lin et al., "Achromatic metalens array for full-colour light-field imaging," *Nat. Nanotechnol.* **14**(3), 227–231 (2019).

97. S.-H. Baek et al., "Compact single-shot hyperspectral imaging using a prism," *ACM Trans. Graphics* **36**(6), 217 (2017).

98. B. H. Chen et al., "GAN metalens for pixel-level full-color routing at visible light," *Nano Lett.* **17**(10), 6345–6352 (2017).

99. M. Chen et al., "Full-color nanorouter for high-resolution imaging," *Nanoscale* **13**(30), 13024–13029 (2021).

100. Z. Jin et al., "Complex inverse design of metaoptics by segmented hierarchical evolutionary algorithm," *ACS Nano* **13**(1), 821–829 (2019).

101. G. Strang, "Computational science and engineering," *Optimization* **551**(563), 571–586 (2007).

102. C. Li et al., "An efficient augmented Lagrangian method with applications to total variation minimization," *Comput. Optim. Appl.* **56**, 507–530 (2013).

103. J. Yang and Y. Zhang, "Alternating direction algorithms for l_1-problems in compressive sensing," *SIAM J. Sci. Comput.* **33**(1), 250–278 (2011).

104. S. Zhuo and T. Sim, "Defocus map estimation from a single image," *Pattern Recognit.* **44**(9), 1852–1858 (2011).

105. J. Salvi et al., "A state-of-the-art in structured light patterns for surface profilometry," *Pattern Recognit.* **43**(8), 2666–2680 (2010).

106. J. Geng, "Structured-light 3D surface imaging: a tutorial," *Adv. Opt. Photonics* **3**(2), 128–160 (2011).

107. Y. Wu and H. Shroff, "Faster, sharper, and deeper: structured illumination microscopy for biological imaging," *Nat. Methods* **15**(12), 1011–1019 (2018).

108. Phys.org., 2015, "IMEC develops snapshot hyperspectral image sensors," https://phys.org/news/2015-02-imec-snapshot-hyperspectral-image-sensors.html

109. X. Hu et al., "Convolutional sparse coding for RGB + NIR imaging," *IEEE Trans. Image Process.* **27**(4), 1611–1625 (2017).

110. Sony Semiconductor Solutions Corporation, "Polarsens-polarization image sensors," https://www.sony-semicon.com/en/technology/industry/polarsens.html.

111. B. K. Gunturk et al., "Demosaicking: color filter array interpolation," *IEEE Signal Process Mag.* **22**(1), 44–54 (2005).

112. S. Balakrishnama and A. Ganapathiraju, "Linear discriminant analysis: a brief tutorial," *Inst. Signal Inf. Process.* **18**, 1–8 (1998).

113. L. K. Saul and S. T. Roweis, "An introduction to locally linear embedding," 2000, http://www.cs.toronto.edu/~roweis/lle/publications.html.

114. D. Bank, N. Koenigstein, and R. Giryes, "Autoencoders," in *Machine Learning for Data Science Handbook: Data Mining and Knowledge Discovery Handbook*, L. Rokach, O. Maimon, and E. Shmueli, Eds., pp. 353–374, Springer (2023).

115. M. S. Asif et al., "FlatCam: thin, bare-sensor cameras using coded aperture and computation," *IEEE Trans. Comput. Imaging* **3**(3), 384–397 (2017).

116. J. K. Adams et al., "Single-frame 3D fluorescence microscopy with ultraminiature lensless FlatScope," *Sci. Adv.* **3**(12), e1701548 (2017).

117. A. Ozcan and E. McLeod, "Lensless imaging and sensing," *Ann. Rev. Biomed. Eng.* **18**, 77–102 (2016).

118. V. Boominathan et al., "Recent advances in lensless imaging," *Optica* **9**(1), 1–16 (2022).

119. J. W. Goodman, *Introduction to Fourier Optics*, Roberts and Company (2005).

120. N. Yu et al., "Light propagation with phase discontinuities: generalized laws of reflection and refraction," *Science* **334**(6054), 333–337 (2011).

121. N. Yu and F. Capasso, "Flat optics with designer metasurfaces," *Nat. Mater.* **13**(2), 139–150 (2014).

122. S. So et al., "Revisiting the design strategies for metasurfaces: fundamental physics, optimization, and beyond," *Adv. Mater.* **35**, 2206399 (2022).

123. Z. Lin et al., "End-to-end nanophotonic inverse design for imaging and polarimetry," *Nanophotonics* **10**(3), 1177–1187 (2020).

124. M. Mansouree et al., "Multifunctional 2.5D metastructures enabled by adjoint optimization," *Optica* **7**(1), 77–84 (2020).

125. J.-M. Jin, *The Finite Element Method in Electromagnetics*, John Wiley & Sons (2015).

126. A. F. Oskooi et al., "Meep: a flexible free-software package for electromagnetic simulations by the FDTD method," *Comput. Phys. Commun.* **181**(3), 687–702 (2010).

127. W. Shin and S. Fan, "Choice of the perfectly matched layer boundary condition for frequency-domain Maxwell's equations solvers," *J. Comput. Phys.* **231**(8), 3406–3431 (2012).

128. T. J. R. Hughes, *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*, Courier Corporation (2012).

129. M. Makarenko et al., "Robust and scalable flat-optics on flexible substrates via evolutionary neural networks," *Advanced Intelligent Systems* **3**(11), 2100105 (2021).

130. H.-Y. Li et al., "Intelligent electromagnetic sensing with learnable data acquisition and processing," *Patterns* **1**(1), 100006 (2020).

131. L. Zhang, "Fast stereo matching algorithm for intermediate view reconstruction of stereoscopic television images," *IEEE Trans. Circuits Syst. Video Technol.* **16**(10), 1259–1270 (2006).

132. L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation-based noise removal algorithms," *Physica D* **60**(1–4), 259–268 (1992).

133. N. Ahmed, T. Natarajan, and K. R. Rao, "Discrete cosine transform," *IEEE Trans. Comput.* **C-23**(1), 90–93 (1974).

134. S. G. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.* **11**(7), 674–693 (1989).

135. H. Lee et al., "Efficient sparse coding algorithms," in *Adv. Neural Inf. Process. Syst. 19* (2006).

136. X. Lin et al., "Spatial-spectral encoded compressive hyperspectral imaging," *ACM Trans. Graph.* **33**(6), 233 (2014).

137. C. R. Gonzalez and R. E. Woods, *Digital Image Processing*, Prentice Hall (2002).

138. A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.* **2**(1), 183–202 (2009).

139. Z. Wang et al., "Uformer: a general U-shaped transformer for image restoration," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, IEEE, pp. 17683–17693 (2022).

140. Z. Lu et al., "Transformer for single image super-resolution," in *Proc. IEEE/CVF Conf. Comput. Vision and Pattern Recognit.*, IEEE, pp. 457–466 (2022).

141. A. Chakrabarti and T. Zickler, "Statistics of real-world hyperspectral images," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit.*, IEEE, pp. 193–200 (2011).

142. F. Yasuma et al., "Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum," *IEEE Trans. Image Process.* **19**(9), 2241–2253 (2010).

143. I. Choi et al., "High-quality hyperspectral reconstruction using a spectral prior," *ACM Trans. Graph.* **36**(6), 218 (2017).

144. D. H. Foster et al., "Frequency of metamerism in natural scenes," *J. Opt. Soc. Am. A* **23**(10), 2359–2372 (2006).

145. S. Qiu et al., "Polarization demosaicking for monochrome and color polarization focal plane arrays," in *Int. Symp. Vision, Model. and Vis.*, The Eurographics Association (2019).

146. Y. Ba et al., "Deep shape from polarization," *Lect. Notes Comput. Sci.* **1269**, 554–571 (2020).

147. R. Blin et al., "A new multimodal RGB and polarimetric image dataset for road scenes analysis," in *Proc. IEEE Conf. Comput. Vision and Pattern Recognit. Workshops*, pp. 216–217 (2020).

148. A. Fan et al., "Full-Stokes polarization multispectral images of various stereoscopic objects," *Sci. Data* **10**(1), 328 (2023).

149. H. B. Barlow, "Unsupervised learning," *Neural Comput.* **1**(3), 295–311 (1989).

150. Y. Wang et al., "Generalizing from a few examples: a survey on few-shot learning," *ACM Comput. Surv.* **53**(3), 63 (2020).

151. S. Hochreiter, "The vanishing gradient problem during learning recurrent neural nets and problem solutions," *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.* **6**(2), 107–116 (1998).

152. Z. Jia and H. Su, "Information-theoretic local minima characterization and regularization," in *Int. Conf. Mach. Learn.*, pp. 4773–4783 (2020).

153. S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," in *Int. Conf. Mach. Learn.*, pp. 448–456 (2015).

154. Y. Wu et al., "Phasecam3D—learning phase masks for passive single view depth estimation," in *IEEE Int. Conf. Comput. Photogr.*, IEEE, pp. 1–12 (2019).

Xuemei Hu is an associate research fellow at Nanjing University. She received her BS and PhD degrees from the University of Tsinghua in 2013 and 2018. Her current research interests include computational imaging and computer vision.

Ting Xu is a professor at the National Laboratory of Solid-State Microstructures, College of Engineering and Applied Science, Nanjing University, China. He directs an optics laboratory, and his research interests include photonic metamaterials, metasurfaces, and metadevices.

Biographies of the other authors are not available.