

REVIEW ARTICLE

Open Access

Physical twinning for joint encoding-decoding optimization in computational optics: a review

Liheng Bian^{1,2✉}, Xinrui Zhan¹, Rong Yan¹, Xuyang Chang¹, Hua Huang^{3✉} and Jun Zhang^{1✉}

Abstract

Computational optics introduces computation into optics and consequently helps overcome traditional optical limitations such as low sensing dimension, low light throughput, low resolution, and so on. The combination of optical encoding and computational decoding offers enhanced imaging and sensing capabilities with diverse applications in biomedicine, astronomy, agriculture, etc. With the great advance of artificial intelligence in the last decade, deep learning has further boosted computational optics with higher precision and efficiency. Recently, there developed an end-to-end joint optimization technique that digitally twins optical encoding to neural network layers, and then facilitates simultaneous optimization with the decoding process. This framework offers effective performance enhancement over conventional techniques. However, the reverse physical twinning from optimized encoding parameters to practical modulation elements faces a serious challenge, due to the discrepant gap in such as bit depth, numerical range, and stability. In this regard, this review explores various optical modulation elements across spatial, phase, and spectral dimensions in the digital twin model for joint encoding-decoding optimization. Our analysis offers constructive guidance for finding the most appropriate modulation element in diverse imaging and sensing tasks concerning various requirements of precision, speed, and robustness. The review may help tackle the above twinning challenge and pave the way for next-generation computational optics.

Introduction

Computational optics, which integrates optics with computation, stands out as a powerful technique for high-dimensional optical information acquisition¹. In contrast to traditional optical methods that primarily address the human visual perception of “what you see is what you get”, computational optics first employs diverse modulation elements to couple high-dimensional information (such as spatial, spectral, and semantic dimensions) into a low-dimensional optical field that can be directly measured by existing detectors, referred to as the encoding procedure. Then, such techniques employ algorithms to recover high-dimensional information from the measurements,

referred to as the decoding procedure². Benefiting from the first-encoding-then-decoding mechanism, computational optics surpass traditional optical limitations of low sensing dimension, low light throughput, low resolution, and so on, developed into a significant and competitive field for optical acquisition and reconstruction³. Nowadays, it holds significant value across various fields such as biomedicine, agriculture, and intelligent manufacturing due to its superior and promising performance, even under extreme conditions^{4,5}.

With the great advance of artificial intelligence in the last decade, deep learning has further boosted computational optics with higher precision and efficiency. The deep learning technique dates back to the middle of the 20th century, with theoretical proposals of concepts including backpropagation^{6,7}, and single-layer perceptrons^{8–11}, as shown in Fig. 1. Over the last decades, rapid advancements in parallel computation and big data processing enabled its practical realizations and propelled the second wave of artificial intelligence¹². The advancements in artificial intelligence have also led to the emergence of

Correspondence: Liheng Bian (bian@bit.edu.cn) or Hua Huang (huahuang@bnu.edu.cn) or Jun Zhang (zhjun@bit.edu.cn)

¹State Key Laboratory of CNS/ATM & MIT Key Laboratory of Complex-field Intelligent Sensing, Beijing Institute of Technology, Beijing & Zhuhai, China

²Yangtze Delta Region Academy of Beijing Institute of Technology (Jiaxing), Jiaxing, China

Full list of author information is available at the end of the article

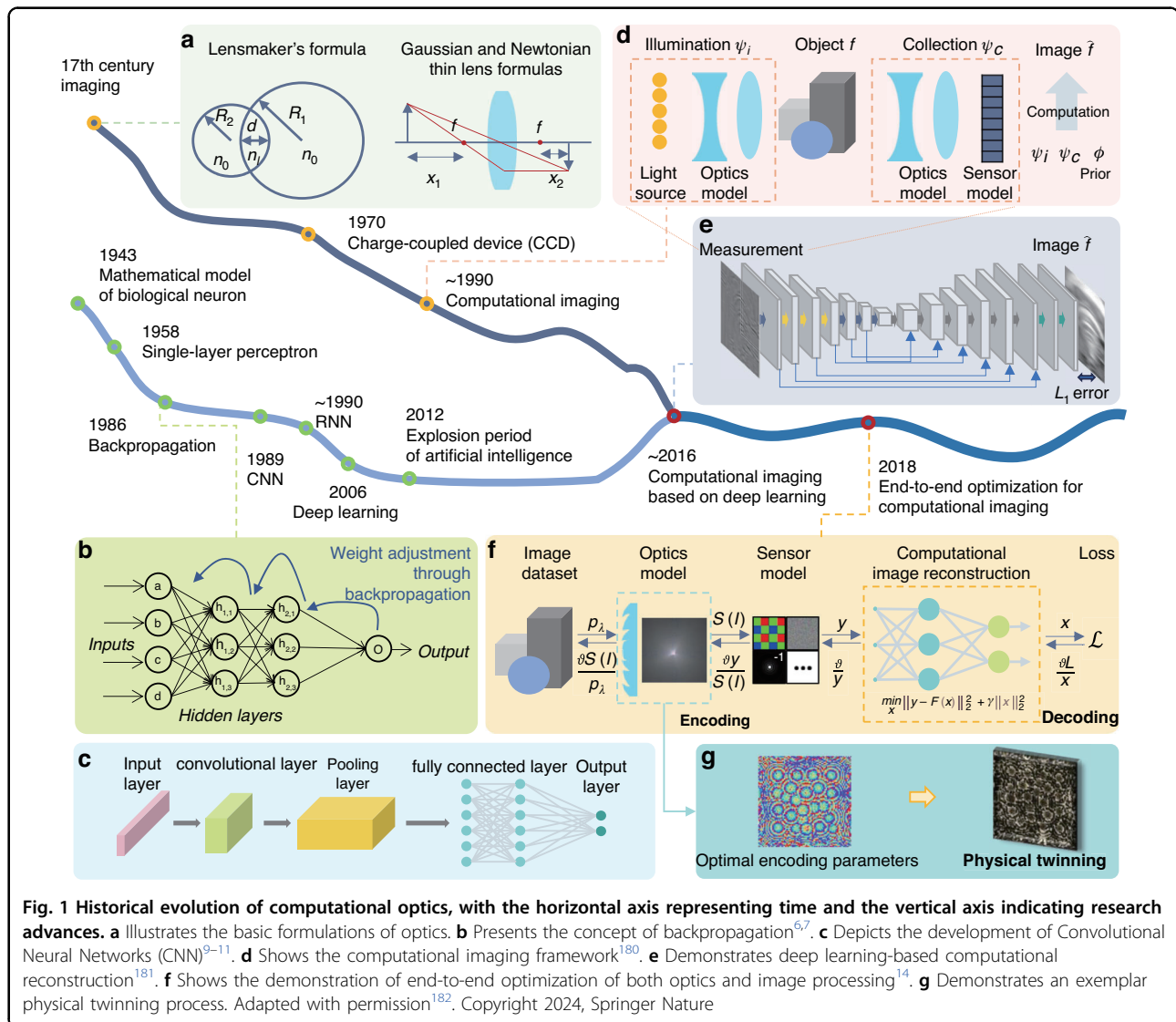
These authors contributed equally: Liheng Bian, Xinrui Zhan

© The Author(s) 2025



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.





deep learning-based computational optics as a pivotal technique. As depicted in Fig. 1(e), in its framework, the deep learning network functions as the decoding process that inputs the acquired encoded measurement and accomplishes image reconstruction or semantic sensing. This approach overcomes limitations inherent in conventional model-based iterative algorithms, particularly in handling complex tasks such as large-scale imaging¹ and real-time sensing¹³. Consequently, it has been widely adopted in various applications such as super-resolution imaging¹⁴, phase imaging^{15,16}, and hyperspectral imaging¹⁷. To further promote efficiency, deep learning-based computational optics focuses on designing high-performance encoding patterns for feature extraction and adopting advanced neural networks for high-quality decoding. However, those methods treat the encoding and decoding processes as separate steps and optimize them

independently, which may not be the optimal choice from the perspective of the whole in most cases.

In most recent studies, an end-to-end optimization method, as illustrated in Fig. 1(f), aims to further enhance the imaging and sensing performance by integrating both encoding and decoding procedures. As shown in Fig. 2, the end-to-end optimization refers to the approaches that utilize an integrated network to optimize both physical encoding and computational decoding parameters. The encoding parameters are then physically twinned in optical systems to achieve high-fidelity imaging and sensing. The system designed by Sitzmann et al. in 2018 is recognized as the first end-to-end optimization demonstration^{14,18}. In this framework, a network model maps the targeted optical field directly to the reconstruction result. The encoding procedure, which accounts for optical models including diffractive light propagation,

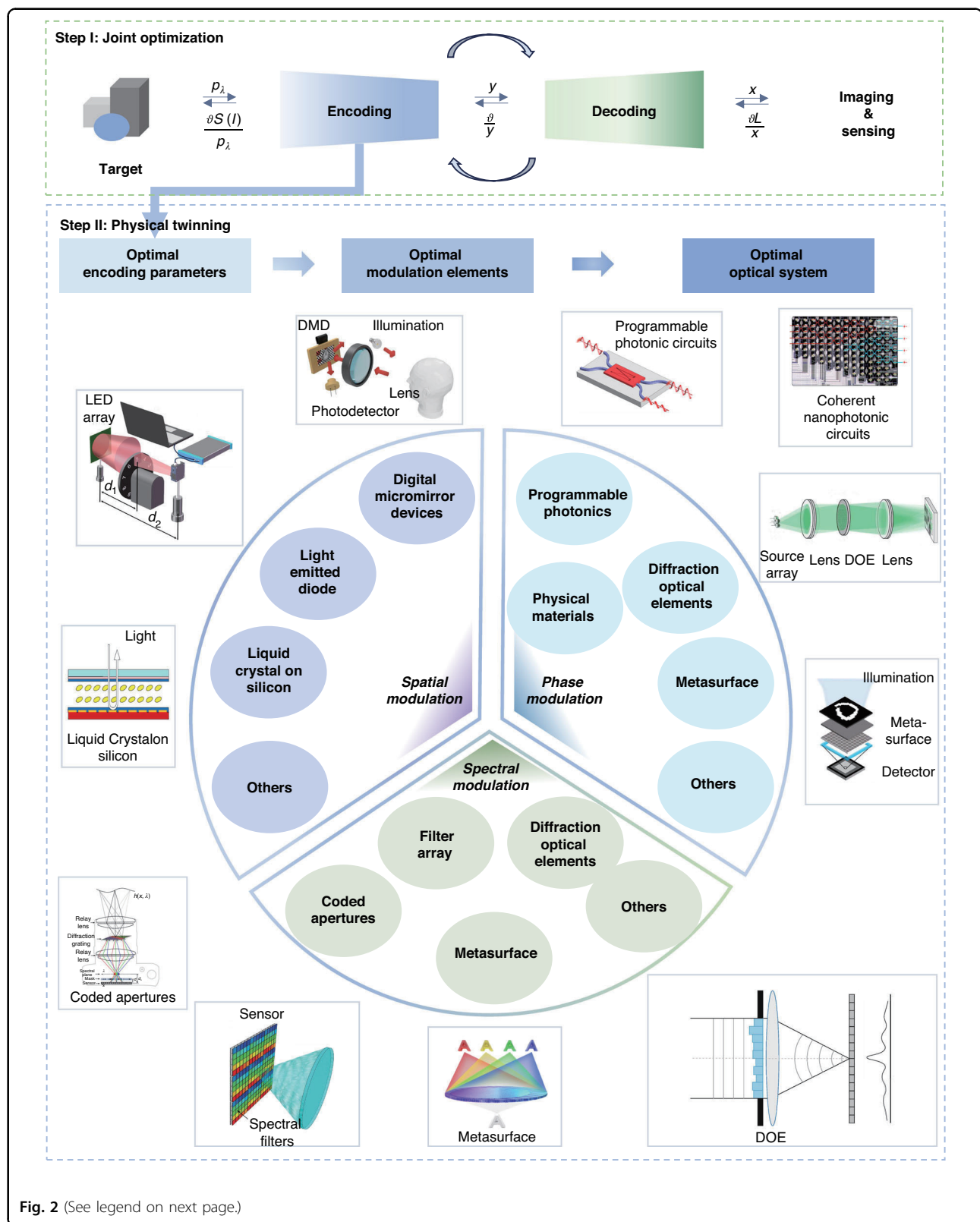


Fig. 2 (See legend on next page.)

(see figure on previous page)

Fig. 2 Two steps for joint optimization in computational optics. First, the encoding and decoding processes are optimized in an end-to-end manner, with the encoding treated as network operations using modulation matrices as parameters and decoding managed by a network for target reconstruction or semantic sensing. Second, the optimized encoding parameters are applied to physical modulation elements, and the acquired measurements are fed into the decoding network for subsequent imaging and sensing. The subfigures present exemplar modulation elements, including Deformable Micromirror Device (DMD, adapted with permission²⁵. Copyright 2018, Springer Nature), liquid crystal (LC, adapted with permission¹⁸³ under Creative Common CC BY license. Copyright 2018, Huang et al., and LED array (adapted with permission¹⁸⁴. Copyright 2024, Elsevier) for amplitude modulation; Diffractive Optical Element (DOE, adapted with permission³¹ under Creative Common CC BY license. Copyright 2023, Bernstein et al.), metasurface (adapted with permission³⁰ under Creative Common CC BY license. Copyright 2022 Springer Nature), phase-only Spatial Light Modulator (SLM) and others, such as programmable photonic circuit (adapted with permission²⁹. Copyright 2020, Springer Nature) and nanophotonic circuit (reprinted with permission²⁷. Copyright 2017, Springer Nature) for phase modulation; and filter array (adapted with permission³² under Creative Common CC BY license. Copyright 2013, Hagen et al.), DOE (adapted with permission¹²⁵ under Creative Common CC BY license. Copyright 2022, Springer Nature), metasurface (adapted with permission³⁵ under Creative Common CC BY license. Copyright 2022, Springer Nature), SLM and others, such as coded aperture (adapted with permission³³. Copyright 2014, Association for Computing Machinery) for spectral modulation

depth, and wavelength-dependent effects, is digitally twinned to neural network layers. The decoding procedure, represented by a neural network, integrates the encoding procedure within its network layers. Through iterative training of the encoding-decoding network, the decoding loss function provides feedback for simultaneously optimizing the encoding parameters. The optimized encoding parameters assist in improving the decoding results, thereby leading to a notable efficiency enhancement. The end-to-end optimization strategy demonstrates effective improvement on encoding efficiency and decoding precision and has been favored in real-time, high-precision imaging and sensing^{13,19}.

With the optimized encoding parameters, there follows a subsequent reverse physical implementation in the end-to-end optimization framework. It requires modulation elements to physically realize the optimized encoding parameters, which we refer to as the “physical twinning” procedure that denotes the mapping process from mathematical design to physical implementation (as shown in Fig. 1(g)). However, the end-to-end optimized encoding-decoding framework struggles to account for all physical restrictions, including bit depth, numerical range, stability, *etc.* Recent studies have attempted to tackle these challenges. First, to deal with the bit depth gap, mathematical binarization algorithms^{20,21}, such as dithering, and binary encoding network operations²², can be applied to translate encoding parameters onto precision-limited devices. For the challenge of numerical range, where optimized modulation matrices contain negative or excessively large values, these matrices can be decomposed into multiple non-negative sub-matrices²¹. Physical twinning is then realized by applying temporal weighting and accumulating measurements from each sub-matrix. Third, environmental factors such as temperature²³ and service life, affect the stability of modulation elements and should also be incorporated as part of the encoder’s optimized parameters. However, accounting for all these variables in the optimization process would lead to a

highly complex and unsystematic structure. Therefore, despite the existing advancements and auxiliary strategies for physical twinning, it remains challenging to balance high-precision results with suitable modulation elements of various experimental requirements.

In this regard, we delve into the physical twinning process in end-to-end jointly-optimized computational optics. As illustrated in Fig. 2, the following sections present common modulation elements across amplitude^{24–26}, phase^{27–31}, and spectral^{32–35} modulation dimensions. In each section, we provide a list of representative modulation elements suitable for joint optimization, along with introductions to their principles, applications, and limitations. This work offers constructive guidance for finding the most appropriate modulation element in diverse tasks concerning various requirements of precision, speed, and robustness, and helps to promote the development of jointly-optimized computational optics.

Amplitude modulation

Amplitude refers to the maximum deviation of light’s electromagnetic field from its average value during wave propagation. The amplitude modulation, also known as intensity modulation, is typically a two-dimensional spatial function that alters light intensity in a controlled manner. This kind of modulation can be achieved in either transmissive or reflective modes, where the light amplitude is controlled by adjusting the local transmission or reflection coefficients of the optical medium³⁶. A common form is binary amplitude modulation, where the amplitude is adjusted between two distinct reflection or transmission values.

Through the last several decades, computational optics systems utilizing amplitude modulation have been widely applied across various fields, such as multi-scale gigapixel photography³⁷, single-pixel imaging²⁵, wavefront shaping³⁸, imaging through turbid media³⁹, ultrasound high-speed imaging⁴⁰, and so on. Multiscale gigapixel photography³⁷ aims to establish a camera array for scalable



gigapixel imaging. Single-pixel imaging, also known as computational ghost imaging, maximizes spatial redundancy by spatial encoding and coupling⁴¹. It applies a sequence of spatial modulation patterns and uses a single-pixel detector to record the cumulative light intensity^{19,22}. Other systems based on amplitude modulation, such as wavefront shaping that spatially modulates incident light to focus or penetrate scattering media³⁸, also adopt a similar imaging model. In these systems, amplitude modulation can be mathematically represented as a Hadamard product during the encoding process, while the modulated light is fed into the decoding algorithms to achieve specific goals.

Adler et al. introduced end-to-end joint optimization into amplitude modulation for compressed imaging^{13,42}. When applied to the MNIST dataset⁴³, their method achieved an impressive 6.46% classification error, significantly outperforming the state-of-the-art error rate of 41.06% at a sampling ratio of 0.1. Subsequent studies have further explored and validated joint optimization in amplitude modulation, leading to wide applications in real-time imaging⁴⁴, object tracking⁴⁵, and terahertz imaging⁴⁶.

Amplitude modulation is mainly implemented using Spatial Light Modulators (SLMs), which can spatially manipulate electromagnetic waves. In the following of this section, we introduce three widely used SLMs for amplitude modulation, including Deformable Micromirror Device (DMD), Liquid Crystal (LC) devices, and Light Emitting Diode (LED), as shown in Fig. 3. Overall, DMD is the most popular amplitude modulation element due to its high resolution and high-speed modulation capability, though the high-speed modulation is limited to the binary mode only. LC offers a low-cost amplitude modulation solution, but it is insufficient in modulation speed and sensitive to temperature conditions. LED takes advantage of high-speed modulation and color performance. However, its spatial resolution is relatively lower than the others.

Deformable micromirror device (DMD)

DMD was early developed by Hornbeck in the 1980s⁴⁷, starting as a 128×128 array addressed by an underlying array of MOS transistors. Each micromirror in the array can be individually rotated by $+12$ degrees or -12 degrees, representing on and off states for light transmission, respectively. At present, DMD is one of the most popular SLMs for amplitude modulation, mainly due to its high resolution (1920×1080 pixels or higher), high refresh rate for binary patterns (22 kHz or higher). Consequently, DMD-based computational optical systems benefit from the high-speed modulation for reliable detection at even invisible wavelengths such as near-infrared⁴⁸ range and mid-infrared⁴⁹ range.

DMD-based computational imaging setups are widely adopted in various experimental end-to-end joint

optimization systems, with convolutional neural networks^{46,50}, recurrent neural networks⁵¹, etc. These systems have demonstrated wide applications including real-time single-pixel imaging^{44,52}, super-resolution imaging⁵³, Light Detection and Ranging (LiDAR)⁵⁰, occluded target imaging⁵⁴, imaging through scattering media⁵⁵, terahertz imaging⁴⁶, and so on. To enhance efficiency for semantic sensing tasks, image-free sensing techniques bypass the complex image reconstruction process and extract semantic information directly from compressed measurements^{19,22,56–60}, offering efficient and lightweight sensing solutions. Their physical twinning process using DMD has successfully demonstrated multiple applications, including character classification^{22,58,59}, multi-character recognition¹⁹, segmentation⁵⁶, detection⁵⁷, object tracking⁴⁵, semi-supervised sensing⁶¹, few-shot sensing⁶², optical encryption⁵⁸, optical neural network⁶⁰, and so on. In these experiment systems, end-to-end optimization shows significant improvement on decoding performance. For instance, it enhanced Chinese characters' perception accuracy by up to 24% compared to conventional methods⁵⁵.

Despite the high-speed modulation and effectiveness at invisible wavelengths, DMD has certain limitations. First, each micromirror on DMD is fixed to either an "on" or "off" position, and therefore provides high-speed modulation only for binarized patterns. For grayscale modulation patterns with multiple bits, it needs to be implemented multiple times at an exponential scale. Each grayscale pixel's intensity is temporally weighted within a predefined time frame. Such a process largely degrades modulation efficiency. For example, a common DMD can only generate 8-bit grayscale patterns at a frame rate of around 250 Hz²⁰. To address this limitation, the pattern binarization strategy has been introduced to approximate grayscale modulation matrices into binary ones²⁰. However, the pattern binarization process disrupts the optimal matching between physical encoding and computational decoding, and consequently results in information loss and deviates from global optimum. Fu et al.²² report a two-step training strategy for joint optimization. In the first step, an end-to-end network is trained with a global scaling factor α relating the optimal modulation pattern values and their corresponding binary ones. In each backward propagation, the grayscale weights are updated using the gradients, and the optimal binary weights can be obtained at the end of the first training stage. In the second step, the encoding parameters are fixed as binary values, and the decoding network is fine-tuned to better match the fixed binarized encoder and improve sensing performance. This two-step training strategy can effectively enhance optical performance with high-speed binary modulation. Another limitation of DMD is that it requires a fine-tuned lens-based imaging system to

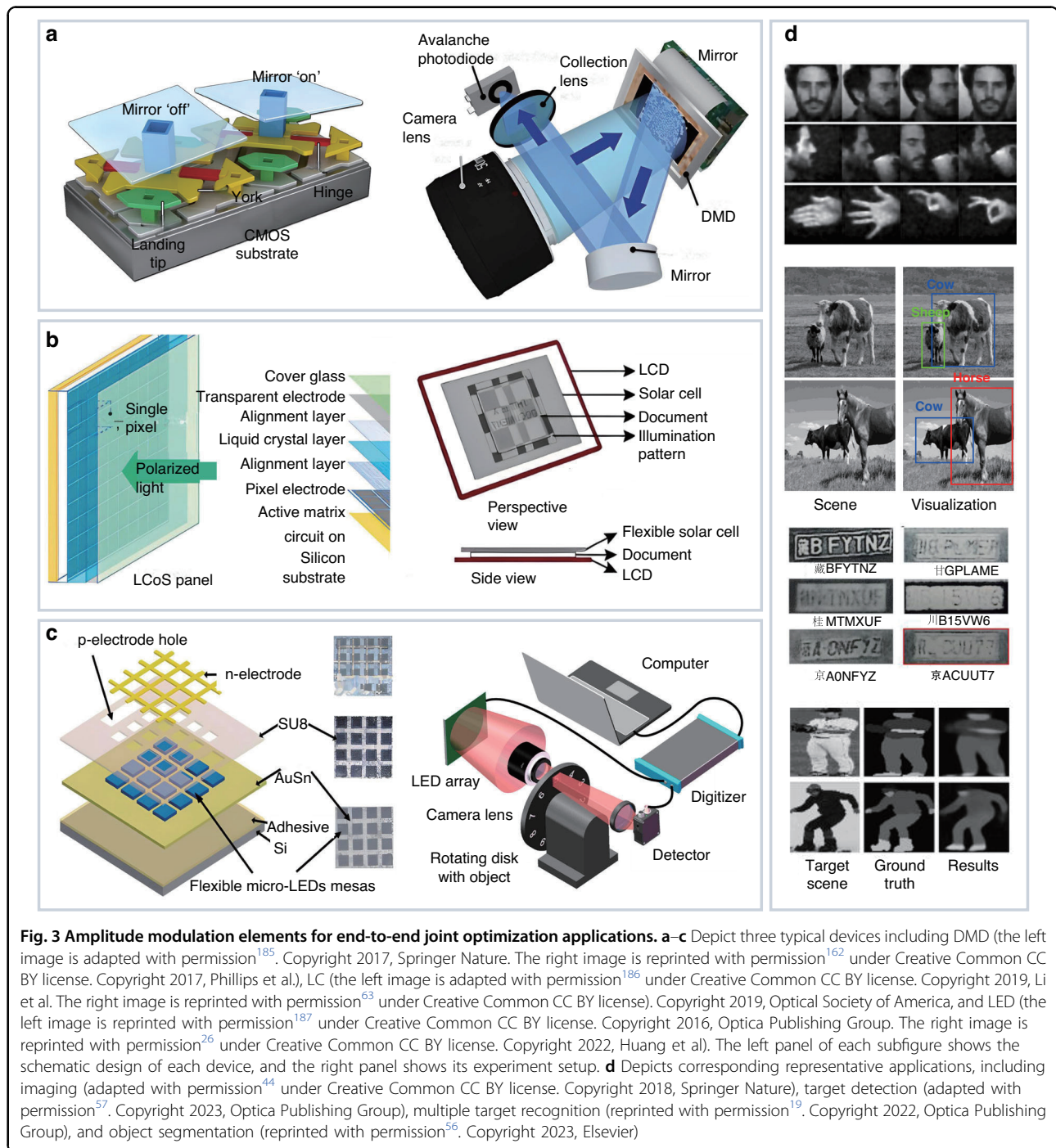


Fig. 3 Amplitude modulation elements for end-to-end joint optimization applications. **a–c** Depict three typical devices including DMD (the left image is adapted with permission¹⁸⁵. Copyright 2017, Springer Nature. The right image is reprinted with permission¹⁶² under Creative Common CC BY license. Copyright 2017, Phillips et al.), LC (the left image is adapted with permission¹⁸⁶ under Creative Common CC BY license. Copyright 2019, Li et al. The right image is reprinted with permission⁶³ under Creative Common CC BY license). Copyright 2019, Optical Society of America, and LED (the left image is reprinted with permission¹⁸⁷ under Creative Common CC BY license. Copyright 2016, Optica Publishing Group. The right image is reprinted with permission²⁶ under Creative Common CC BY license. Copyright 2022, Huang et al). The left panel of each subfigure shows the schematic design of each device, and the right panel shows its experiment setup. **d** Depicts corresponding representative applications, including imaging (adapted with permission⁴⁴ under Creative Common CC BY license. Copyright 2018, Springer Nature), target detection (adapted with permission⁵⁷. Copyright 2023, Optica Publishing Group), multiple target recognition (reprinted with permission¹⁹. Copyright 2022, Optica Publishing Group), and object segmentation (reprinted with permission⁵⁶. Copyright 2023, Elsevier)

project structured light patterns. This requirement is easily corrupted by aberrations and poses challenges on establishing compact imaging systems⁶³.

Liquid crystal (LC)

Liquid crystal (LC) materials are organic compounds that exhibit both liquid and crystalline properties. They are characterized by anisotropic dielectric constants and

refractive indices. By leveraging these anisotropic properties, the orientation of liquid crystal molecules can be controlled via an electric field, allowing for the modulation of light propagation direction. This principle underlies the development of Liquid Crystal Display (LCD), which utilizes the birefringence of liquid crystal molecules. As a widely used SLM, LCD achieves grayscale control by applying different voltages to pixel units in conjunction

with orthogonal polarizers. Over time, advancements led to the development of the Liquid Crystal on Silicone (LCoS) technique, which combines silicon-based integrated circuits to offer a higher spatial resolution (up to million pixels), better light energy utilization (70–80% light throughput efficiency, and 2000 lumens/cm² luminous densities⁶⁴), and faster pixel switching speed (response time of ~ 1 ms at gray level⁶⁴) compared to LCD⁶⁵. LC-based modulation elements maintain certain advantages over DMD, including the ability to project grayscale patterns, the applicability of phase modulation, and the elimination of lens-related degradation such as optical distortion, aberrations, and chromatic dispersion⁶³.

LCD and LCoS have been widely employed for amplitude modulation benefiting from their advantages in high-precision modulation and constructing lensless imaging systems. For instance, Zhang et al. adopted LCD to construct a compact lensless single-pixel imaging system in which the compact scanner is as thin as 2.48 mm, and achieved grayscale and true-color scanning as well as accurate optical character recognition⁶³. LCoS has been widely applied in Structured Illumination Microscopy (SIM), including 2D SIM⁶⁶ and 3D SIM⁶⁷, offering high frame rates with moderate cost. For instance, Lu-Walther et al. designed a FLCoS-based two-beam SIM fluorescence microscope and realized a lateral resolution of ~100 nm at a maximum raw data acquisition rate of 162 frames per second (fps) with a region of interest (ROI) of 16.5 × 16.5 μm²⁶⁸. Hannebelle et al. designed a SIM system that places LCoS at the primary image plane of microscope⁶⁹. This system, in combination with a polarizing beam splitter (PBS) in the amplitude mode, serves a computer-controlled modulation function that reduces demands for precise positioning and alignment of multiple optical elements. The striped illumination pattern of SIM⁷⁰ can also be jointly optimized under certain constraints in the encoder module. Therefore, although these LCoS-based imaging systems do not currently utilize end-to-end joint optimization, their flexible design makes them promising for jointly-optimized physical twinning.

LC-based amplitude modulation elements have certain limitations. First, constrained by the intrinsic properties of LC, these elements are sensitive to temperature variations, leading to performance degradation at extremely high and low temperatures. For instance, the operating temperature of the common Thorlabs WUXGA LCoS is restricted between 10 °C and 40 °C²³. Second, the modulation speed of LC-based elements still lags behind that of DMD, for around 100 Hz⁶³. Ferroelectric Liquid Crystals on Silicon (FLCOS) maintain a faster modulation speed of up to several kHz, more than an order of magnitude higher than nematic liquid crystals⁷¹, which is however relatively more complex to fabricate and may suffer from significant degradation over long-term use or frequent switching^{72,73}.

Light emitted diode (LED)

The LED array consists of a two-dimensional diode matrix where cathodes are connected by rows and anodes by columns. By controlling the electrical current through each diode to vary the intensity of the emitted light, it enables precise modulation of light's amplitude. Within an LED array, the individual current of each diode in each column or row can be independently controlled, facilitating end-to-end optimized modulation patterns for physical twinning. Moreover, LED can achieve remarkably high frequencies for amplitude modulation, with modulation speed up to around 50MHz⁷⁴.

LED has diverse applications in computational imaging as a structured light encoding device. For instance, Xu et al. report a computational ghost imaging scheme utilizing an LED-based high-speed illumination module. This approach achieved a frame rate of 1000 fps with a resolution of 32 × 32 pixels⁷⁵. Salvador et al. used a low-cost color LED array for structured illumination, enabling a high refresh rate of up to 10 kHz. This system can be further extended for 3D imaging by incorporating additional photodiodes placed at different positions⁷⁶. Although the above-mentioned publications have not realized physical twinning of end-to-end optimization, the flexible design of LED offers the potential to realize learning-based encoding patterns.

LED faces a serious limitation of low spatial resolution. The size of a typical LED array with a resolution of 32 × 32 pixels is around 64 mm × 64 mm⁷⁵. The limited spatial modulation precision limits the system's ability to produce fine modulation patterns. To achieve a higher spatial resolution, it is necessary to construct an expanded LED array. However, an extended LED array is too large in size for compact imaging systems. Moreover, it becomes challenging for a single FPGA to provide enough I/O ports for a large-scale LED array⁷⁵, leading to a more complex system with a slower response speed. This increase in complexity can complicate the setup and potentially reduce the efficiency and practicality of the system for certain applications.

Others

In addition to the above common SLMs, there are alternative elements for amplitude modulation. For example, Hahamovich et al. implemented cyclic patterns using a spinning mask, achieving rapid amplitude modulation of up to 2.4 MHz, surpassing the speed limitation of configurable SLMs and enabling real-time visualization of dynamic objects⁷⁷. However, this technique requires a fixed design of cyclic patterns, making it not feasible for controllable physical twinning in end-to-end joint optimization. There are also several other materials with potential for joint optimization. For example, He et al. realized single-pixel neutron imaging using Hadamard



masks made of gadolinium oxide (Gd_2O_3), which had a considerable resonance and absorption cross section for thermal neutrons⁷⁸.

Phase modulation

Both amplitude and phase are fundamental dimensions for describing the wave characteristics of light⁷⁹. Phase represents the propagation delay caused by the interaction between incident wavefront and matter. Phase modulation alters the spatial phase distribution of a wavefront, thereby influencing light convergence, propagation, and delay to achieve specific objectives. The phase modulation-based encoding layer typically involves a Hadamard product between the incident wavefront and the phase pattern. In experimental implementation, this process occurs automatically as the light propagates through the patterns. Compared to amplitude modulation, phase modulation preserves a higher ratio of light intensity.

Phase modulation has been widely applied in computational optics. In quantitative phase imaging^{80–82}, phase modulation is able to extend imaging throughput, field of view, and imaging resolution. Adaptive optics⁸³ employs phase modulation to correct wavefront distortions caused by atmospheric turbulence or optical aberrations, significantly enhancing image resolution in astronomy and microscopy. Wavefront shaping⁸⁴ uses phase modulation to control light propagation through complex media, enabling deep-tissue imaging. Recent years have seen a surge in end-to-end optimization of the phase modulation patterns and imaging or sensing processes¹⁴. A representative example is diffractive deep neural Networks (D^2NN)⁸⁵, as shown in Fig. 4(f). D^2NN consists of several diffraction layers that represent optical operation, each of which is composed of hyperparameters that represent physical light-transferring procedures. The implementation of D^2NN usually requires two steps. First, in the training stage, a task-based D^2NN is built and trained until convergence, with the trainable hyperparameters denoting phase modulation patterns. In the second step, the inference stage, the reconstructed images or sensing results can be directly visualized at the detector when light propagates through these phase patterns. This method enables D^2NN to achieve high-precision inference^{86,87} at a speed of light propagation without external power, potentially overcoming the limitations of high energy consumption and low efficiency in electronic computing.

The above-mentioned studies were based on various phase modulation elements. As shown in Fig. 4, typical phase modulators include Diffraction Optical Element (DOE), phase-only Spatial Light Modulator (SLM), and metasurface. Each type of modulator has its own advantages and disadvantages. DOE provides low-cost and

plug-and-play phase modulation, but it is usually bulky and cumbersome. Phase-only SLM offers high spatial resolution and programmable modulation using liquid crystal, but it is relatively expensive and bulky and has low optical efficiency. Metasurface is compact and lightweight, allowing for easy integration, but it is difficult to fabricate and involves high design costs.

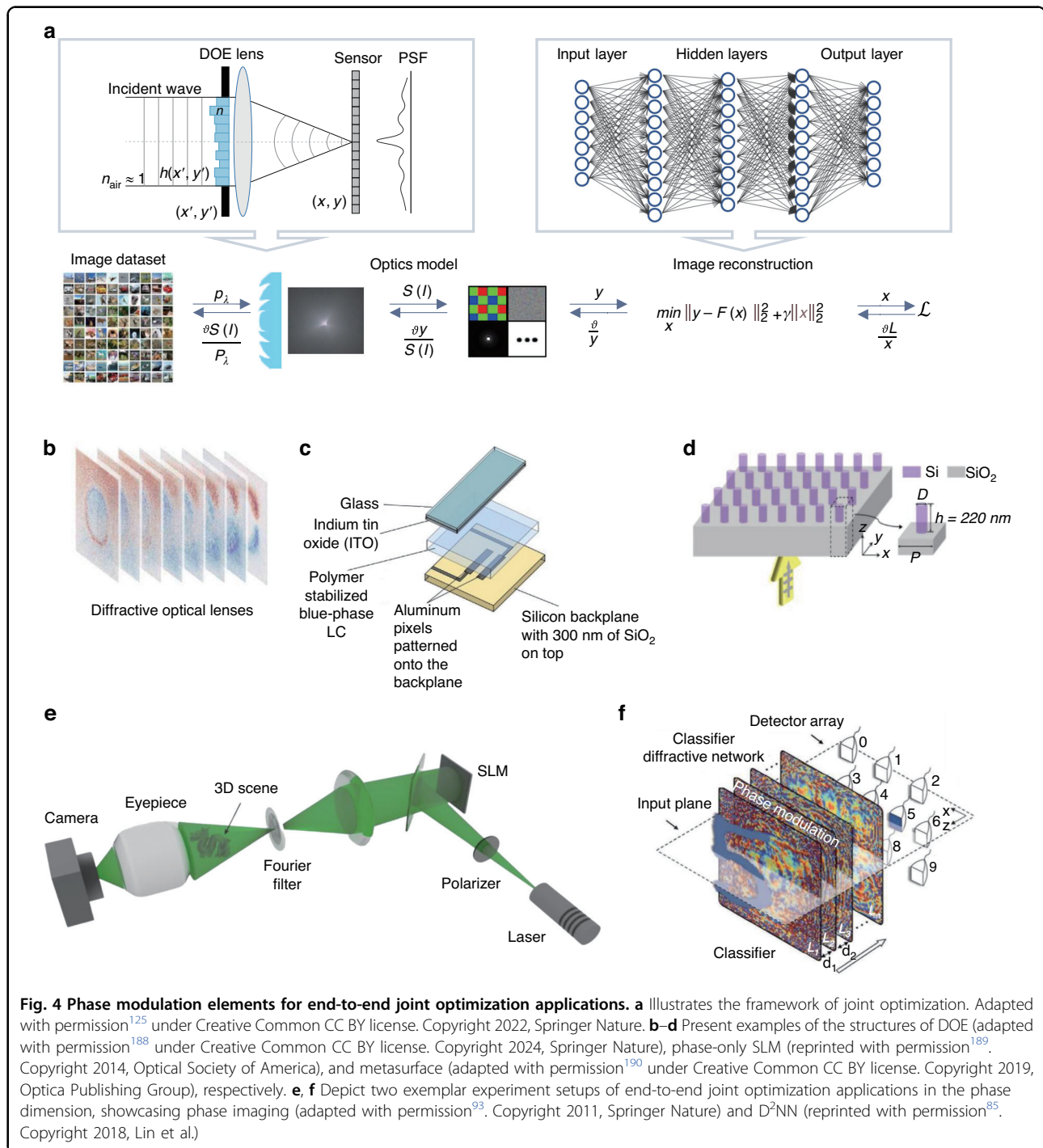
Diffraction optical element (DOE)

Diffraction Optical Element (DOE) manipulates the phase of light by introducing optical path differences on grating's structure. These elements can precisely control the diffraction angles, which affect the direction and distribution of the diffracted light. Designing DOEs involves prior calculations to determine the optimal arrangement and dimensions of the microstructures. The techniques such as rigorous coupled-wave analysis⁸⁸ and finite-difference time-domain (FDTD) simulation⁸⁹ are commonly employed to model and optimize DOE performance. The fabrication of DOEs typically utilizes high-precision lithography or 3D printing techniques to pattern microscopic structures on silicon or glass substrates⁹⁰.

DOEs have been employed in numerous applications such as beam shapers, beam splitters, and diffusers⁹¹. Due to their low cost (as low as a few dozen dollars) and plug-and-play convenience, DOEs have naturally become a common choice for physical twinning in end-to-end joint optimization. A notable study by Li et al. introduced a joint optimization approach that optimizes both the DOE placed in front of the projector lens and the compensation network used for image deblurring⁹². This technique can extend the projector's Depth-of-Field (DOF) by a minor modification to the projector. The compensation network demonstrated significant advancement in radiometric compensation, particularly in terms of computational efficiency and image quality. Shi et al. demonstrated optimized 3D printing for fabricating diffractive layers, showing great performance of real-time 3D holography with low memory requirement (below 620 KB), high frame rate (60 Hz), and high resolution (1920×1080 pixels) on a single consumer-grade graphics processing unit⁹³. To address the discrepancies between optimized digital modulation patterns and the actual fabricated ones resulting from manufacturing imperfections and model approximations, decoding networks were typically fine-tuned through several iterations using the real calibrated patterns.

Although DOE is highly effective for phase modulation, it has several inherent limitations that impede its broader applications. First, the fixed structural configuration limits DOE's adaptability in dynamic optical systems⁹⁴. This rigidity means that once a DOE is fabricated, its phase modulation properties cannot be altered without creating an entirely new element, making real-time adjustment





impossible. Second, DOE usually suffers from a low diffraction efficiency (40% to 70% for binary DOEs)⁹⁵, where not most incident light is effectively modulated as desired. This results in low light efficiency and unwanted scattering or diffraction orders, which degrade the performance of optical systems. Third, most DOEs are bulky and cumbersome⁹¹, posing an integration challenge for compact and portable devices.

Phase-only spatial light modulator (SLM)

Phase-only SLM modulates the phase of light in a pixel-wise manner, typically using liquid crystal⁹⁶ or micro-electromechanical systems (MEMS)⁹⁷. These devices can independently adjust the optical properties of each pixel in response to electronic signals, which alter the refractive index (in liquid crystal SLM) or the physical displacement (in MEMS SLM) to achieve desired phase modulation.

SLM is crucial for applications that require dynamic manipulation of phase, including both conventional optical imaging and end-to-end joint-optimized computational imaging⁹⁸. In terms of conventional optical imaging, SLM can dynamically generate and project high-quality holograms to realize holographic imaging⁹⁹. In adaptive optics, SLM is used to correct wavefront distortions in real time, improving image resolution in astronomical observations and microscopy¹⁰⁰. Besides, SLM also plays a crucial role in beam shaping¹⁰¹, where it is used to tailor the intensity profile of laser beams. In terms of end-to-end joint optimized computational imaging, Chen et al. report a large-scale, cost-effective, complex-valued, and reconfigurable diffractive all-optical neural network system in the visible range based on cascaded transmissive twisted nematic SLMs with a distinct statistical adversarial property¹⁰². Wang et al. used SLM to implement D²NN and realized orbital angular momentum spectrum measurement¹⁰³. Pinilla et al. jointly optimized the phase patterns in front of an image sensor and realized achromatic extended-depth-of-field imaging¹⁰⁴. Compared to conventional lens-based imaging systems, they achieved around 2 dB improvement on PSNR. Due to its high accuracy and programmable phase modulation capabilities, the SLM-based encoding layer can accurately implement digital patterns with minimal mismatch caused by fabrication errors.

Despite the dynamic and pixel-wise modulation capabilities, SLM has several limitations. First, high-performance and high-resolution SLM is usually expensive at tens of thousands of dollars¹⁰⁵. Second, SLM also suffers from low diffraction efficiency (60% to 90%) due to light loss at the liquid crystal layers or MEMS mirrors¹⁰⁶. Third, the calibration of SLM is generally complex and time-consuming¹⁰⁷.

Metasurface

Metasurface is artificially designed and manufactured with special physical properties, such as negative refractive index, negative magnetic permeability, and negative dielectric constant¹⁰⁸. By altering the basic unit structure, shape, orientation, arrangement and properties, metasurface can effectively manipulate various dimensions of light waves. In terms of phase modulation, metasurface introduces phase shifts through the interaction of light with the meta-atoms. Each meta-atom can be designed to impart a specific phase delay, which cumulatively shapes the wavefront of the transmitted or reflected light. Benefiting from the advantages of lightweight, easy integration, and high light efficiency¹⁰⁹, metasurface is particularly suitable for physical twinning of phase modulation in end-to-end joint optimization.

Metasurface has been widely applied in various optical imaging/sensing techniques^{15,110}. In conventional optical

imaging, Zheng et al. developed metasurfaces for generating high-resolution, high-efficiency holographic images, achieving a light efficiency of 80%¹¹⁰. Subsequently, various metasurface-based holography techniques have demonstrated great success^{111–113}. For joint optimization, Ghorbani et al. explored the use of metasurface to construct D²NN, which can perform on-chip multi-channel sensing in the visible range³⁰. Tseng et al. developed a fully differentiable learning framework that learns a physical metasurface structure in conjunction with a neural feature-based image reconstruction algorithm, which achieved an order of magnitude lower reconstruction error than existing approaches¹¹⁴. While the above-mentioned passive-structure-based D²NN platforms can perform complex functions based on computer-aided neural network design, they still rely on computers to optimize parameters and require prior environment information. Liu et al. report a programmable metasurface where each meta-atom on the metasurface was integrated with two amplifier chips, acting as an active artificial neuron. This architecture can handle various deep learning tasks including image classification, mobile communication encoding/decoding, and real-time multi-beam focusing¹¹⁵. Similar to DOE-based techniques, the metasurface-based phase modulation encoding process also suffers from mismatches between optimized digital patterns and actual fabricated ones, which would degrade the performance of decoding network. To address this issue, most studies typically refine the decoding network parameters using the real calibrated patterns.

Despite the above advantages, metasurface faces several limitations for phase modulation. First, the fabrication process is complex and costly, requiring high manufacturing precision for designed microscopic structures¹¹⁶. Second, while metasurface offers excellent control over multiple optical properties, it suffers from efficiency loss due to imperfect nanostructure fabrication and alignment¹¹⁷. Third, the scalability of metasurface is limited by the difficulty in maintaining uniformity across large areas¹¹⁸. Further studies of physical twinning and low-cost fabrication techniques continue to be crucial for advancing the practical applications of these advanced optical materials.

Others

Besides the above-mentioned DOE, SLM and metasurface, there are also some other devices available for phase modulation. For example, scattering/holography diffusers, which are widely used for phase modulation and can achieve super-resolution and large depth-of-field tasks, are seldom considered for optimization^{80,119}. Semiconductor optical phased arrays (OPA) offer a fast modulation speed of up to 10 GHz, but their numerical randomness makes reverse design challenging²⁸. DMD



can be employed for phase modulation with specially designed optical circuits^{120,121}. However, DMDs are essentially binary modulators, and their technological maturity and diffraction efficiency are often insufficient for multiple-phase modulation applications¹²². Nanophotonic circuits based on optoelectronic devices offer programmability but are hindered by a large footprint, posing challenges for the realization of large-scale neural networks²⁷. While the aforementioned novel devices can realize phase modulation, their complex structures, randomness, and low diffraction efficiency and resolution hinder end-to-end joint optimization. Thus, it is essential to comprehensively evaluate the use and performance of these devices when selecting a phase modulator.

Spectral modulation

Dating back to the 17th century, spectrum was defined to represent the light intensity of a range of wavelengths¹²³, which reveals the intrinsic spectral frequency of an electromagnetic wave. Spectral modulation denotes selecting and encoding the intensity of a range of wavelengths, namely a function of wavelength. This modulation relies on manipulating different optical parameters in the propagation path, such as the reflection index or transmittance¹²⁴. When the light goes through the medium, the intensity or phase of each wavelength will change differently, leading to a change in spectrum. Generally, there are three common strategies for spectral modulation, including directly using specific materials for filtering, indirectly using diffraction to map spectral information into phase for modulation, and indirectly using refraction to map spectral information into space for amplitude modulation¹²⁵.

The development of spectral modulation advances computational spectral imaging by selective manipulation and precise encoding of different wavelengths of light. Different from the conventional one-dimensional or two-dimensional scanning-based spectral imaging techniques^{32,126}, computational spectral imaging allows for the acquisition of two-dimensional modulated projection data of the target spatial and spectral information with a single imaging exposure^{32,125}. Through subsequent data decoupling and processing, the three-dimensional spectral data cube can be reconstructed¹²⁵. In recent years, there has been a surge in end-to-end optimization of spectral modulation in computational imaging, as shown in Fig. 5. Similar to amplitude and phase modulation, spectral modulation in the encoding layer is represented by a Hadamard operation, with each spectral component integrated through multiple network channels. Considering the physics-based propagation of light and its interaction with optical elements or devices, end-to-end optimization utilizes advanced algorithms to enhance imaging resolution and accuracy¹²⁷. It holds applications

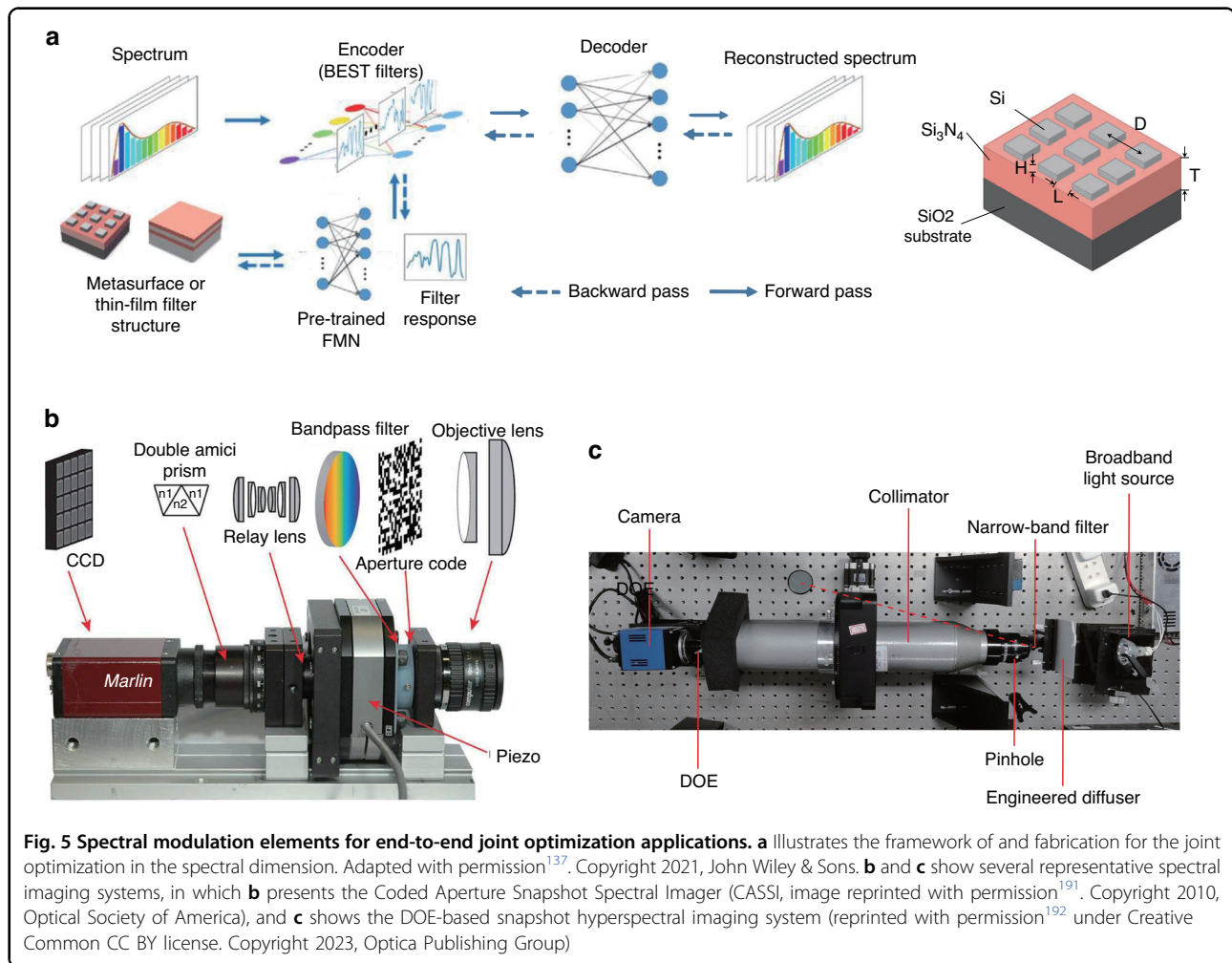
in diverse fields such as astronomical observation, remote sensing, agriculture, and biomedicine^{128,129}.

Conventional spectral modulation elements such as prism and diffraction grating are widely used due to their simplicity. However, they may be constrained by size, diffraction limit, and reliance on light propagation¹³⁰ in practical setups. In recent decades, there have developed multiple devices for spectral modulation, including the following four representative ones of filter array, Diffraction Optical Element (DOE), Spatial Light Modulator (SLM) and metasurface, as shown in Fig.5(b-d). Each one has its respective pros and cons. Filter array offers precise spectral selectivity but maintains spectral crosstalk and limited flexibility¹³¹. DOE can improve light throughput and reduce system volume¹²⁵, but it requires a sophisticated fabrication process and is sensitive to environmental fluctuations. SLM offers high spatial resolution and fast response, but it may suffer from limited light throughput and requires complex drive electronics¹²⁵. Metasurface has versatility and compactness (tens of nm to microns¹³⁰), but it requires precise and costly nanofabrication.

Filter array

Filter array (also known as multispectral filter array) consists of multiple filters arranged in a spatial pattern, and each filter is designed to selectively transmit specific wavelengths of light. By carefully utilizing materials of different optical properties to design arrays, filter array enables precise control over light spectrum at each location. These filters can be designed to allow only a narrow band of wavelengths (narrowband filters) to pass with FWHM being as low as ~ 2 nm^{132,133}, or they can transmit a wide range of wavelengths (broadband filters) or block a certain range (band-stop filters, also known as notch filters). Upon different requirements, they can achieve precise spectral selection and are relatively easy to integrate into different optical devices and systems¹³¹. Additionally, the variety of manufacturing materials, ranging from traditional materials to new ones such as dielectric coatings, glass, or polymers, could achieve diverse results for enhancing or suppressing spectral components.

Benefiting from its spectral selectivity and easy integration advantages, filter array has been extensively applied in computational optics. Studies mainly focused on optimizing the pattern of filter arrays using deep learning methods¹³⁴. For instance, Zhang et al. denoted the transformation matrix of filter array using an encoding neural network, which was learned simultaneously with the decoding network¹³⁵. To realize physical twinning, the network was designed to learn non-negative and smooth spatial-spectral patterns. Based on the joint optimization framework, the results improved PSNR for more than 1.7 dB compared to traditional methods. In recent years,



broadband filters have attracted much attention for hyperspectral imaging due to their high throughput. Zhang et al. designed a broadband-coded hyperspectral camera based on deep learning and achieved a processing speed increase of 7000–11,000 times and approximately a 10-fold improvement on noise tolerance¹³⁶. Using 16 broadband filters, the camera realized multispectral imaging of 480×640 pixels at 1 nm spectral resolution in the range of 400–700 nm. Song et al. reported a spectral encoder and decoder network for joint optimization of filters. Compared to conventional non-learning methods, the designed filter array realized 30 times enhancement of reconstruction accuracy and better tolerance on fabrication errors in the spectral range of 400–700 nm^{137,138}.

Despite its versatility, filter array has certain limitations. One constraint is their finite bandwidth and spectral crosstalk. End-to-end optimization has demonstrated its ability to alleviate this issue¹³⁵. Besides, an existing filter array has fixed spectral characteristics, limiting its adaptability to changing experimental requirements or spectral ranges.

DOE

DOE operates on the diffraction principle to distinguish different wavelengths as light propagates. It can realize complex optical functionalities¹³⁹ with high throughput, enabling high-precision spectral modulation with spectral peak location accuracy better than 1 nm and peak separation resolution of 3 nm¹⁴⁰. Additionally, DOE's compact design ($\sim 1 \mu\text{m}$ thickness¹⁴¹) allows integration into miniaturized optical systems, making them suitable for portable and space-constrained applications.

In recent years, DOE has been widely applied for spectral imaging. The common DOE-based spectral imaging framework builds on the DOE-Fresnel diffraction or DOE-Lens system¹²⁵. For instance, Lv et al. built an Aperture Diffraction Imaging Spectrometer (ADIS), which employs diffraction-based spatial-spectral projection for encoding. It improved the diffraction efficiency to 75% and achieved sub-super-pixel spatial resolution (1056×1536 pixels) in the spectral range of 453–646 nm¹⁴². Meng et al. reported a diffractive optical network-based multispectral imaging system, which

adopted 3D-printed diffractive surfaces. Considering the layer-to-layer misalignments, they were modeled as random variables and incorporated into the forward training model. In this way, the deep learning-based evolution of the diffractive surfaces is enforced to converge to solutions that show resilience against implementation errors. The joint optimized system achieved 16-channel snapshot multispectral imaging in the range of 450–700 nm with > 79% transmission efficiency¹⁴³. Baek et al. reported a DOE-based spectral-depth imaging system, which was jointly optimized by a differentiable simulator and a neural network-based reconstruction method. Via automatic differentiation, the end-to-end network optimizes DOE's PSF at different wavelengths and depths, which can simultaneously couple spectral and depth information into a snapshot. To avoid additional complexity of employing physical fabrication constraints, they optimized the unwrapped phase instead of DOE height. Upon a single shot, the technique can simultaneously realize multispectral imaging with 3840×5760 pixels in the range of 420–660 nm and depth imaging at 0.4–2 m³⁴.

While DOE offers multiple advantages, its fabrication requires complex and precise microfabrication. DOE's performance is sensitive to fabrication imperfections, which limits its scalability and increases production costs. Besides, DOE's efficiency is also affected by diffraction efficiency loss (40–70% for binary DOE⁹⁵), which means that a large portion of incident light is not properly guided to the intended diffraction order. In addition, factors such as incident angle and mechanical stress can also affect DOE's stability and performance. These challenges require careful consideration in the processes of optical design, fabrication, and application.

SLM-based coded aperture

In a coded-aperture spectral modulation system, the spectrum is usually dispersed into the spatial domain by a prism or optical grating. Then, a coded aperture is placed to selectively manipulate the dispersed spectrum. One of the most common and flexible devices to achieve a coded aperture is SLM¹⁴⁴, which can block or unblock target light to modulate the spatial distribution of different wavelengths. This modulation allows precise adjustment and selection of specific wavelengths, thus increasing the flexibility and accuracy of spectral imaging. The key advantages of SLM lie in its high resolution, fast response, and programmability.

Benefiting from SLM's flexibility, Li et al. reported a non-mechanical spatio-spectral modulation method that realized single-pixel multispectral imaging with 128×128 pixels in the 420–720 nm range at 2 Hz¹⁴⁵. In recent studies, the employment of SLM in Coded Aperture Snapshot Spectral Imager (CASSI)¹⁴⁶ has gained much attention. CASSI encodes the hyperspectral datacube into

a single compressed measurement and then employs algorithms to reconstruct the underlying hyperspectral information. In the past five years, a series of end-to-end optimization algorithms have emerged to maximize the information content of compressive measurements. Wang et al. utilized a convolutional neural network to jointly optimize the coded aperture and subsequent reconstruction algorithm. By improving the spectral and spatial correlations, the reconstruction accuracy was improved by about 2 dB compared with other methods¹⁴⁷. Shi et al. reported a self-adaptive encoding method for multispectral imaging, which uses SLM's flexibility to adaptively optimize coded patterns¹⁴⁴. In addition, to ensure flexible physical twinning, they designed an element-wise binarization function to transform the optimized continuous encoding matrix into the binary mask form. The adaptive strategy improves hyperspectral reconstruction accuracy by at least 0.8 dB.

SLM has several limitations for spectral modulation. One is its limited diffraction efficiency (60–90%)¹⁰⁶ and spatial resolution (compared with a 65-megapixel film¹⁴⁸), which degrade the spectral imaging performance. Besides, SLM usually requires precise alignment and correction for different wavelengths, and the system may become bulky (compared with metasurface and DOE) and expensive (tens of thousands of dollars¹⁰⁵), limiting its widespread use in integrated designs and devices. Finally, SLM's stability and reliability are easily affected by environmental fluctuations such as mechanical vibration.

Metasurface

Metasurface is a planar optical element comprising arrays of subwavelength meta-atoms (~ 100 nm thickness)^{130,141}. The structure introduces a sudden change of optical properties on the surface, breaking the dependence of traditional optical elements on the propagation path. This structure can interact with the incident electromagnetic field and significantly manipulate the amplitude, phase, and polarization, enabling indirect, compact, and efficient spectral modulation. By nanofabrication, it exhibits comparable modulation properties to those of conventional optical elements such as diffraction gratings and filters^{109,149}, which provides a novel and compact alternative for spectral modulation.

McClung et al. designed a parallel array of metasystems for spectral modulation and imaging¹⁴⁹. The array consists of metasurface-tuned filters and a metalens doublet, mapping the narrowband spatio-spectral information into different detection units. The 20 mg compact device realized multispectral imaging of $\sim 58,000$ pixels and 20 channels in the 795–980 nm range. Makarenko et al. exploited end-to-end training to design an optimized metasurface layer structure, and achieved multispectral imaging with 512×512 pixels from 400 to 1000 nm (204



bands)¹⁵⁰. Zhang et al. established an end-to-end neural network for the joint optimization of metasurface structure and image decoding. This approach leverages the strong dispersion of metasurface to construct a point spread function that varies from 460 nm to 700 nm (25 channels), improving 800×800 image's reconstruction accuracy with up to 1 dB¹⁵¹.

Despite the above demonstrations, metasurface faces several limitations. The fabrication of metasurface requires ultra-high precision and costly nanofabrication, which limits its scalability and increases production cost¹¹⁶. In addition, the conversion efficiency of metasurface for different wavelengths is different, leading to a certain amount of light energy loss¹⁵¹.

Others

In addition to the above widely adopted spectral modulation elements, there are some other alternatives for spectral modulation. For instance, Zhang et al. reported a handheld snapshot multi-spectral camera (THETA) based on a multiplexed wavelength-dependent encoding film¹⁴⁸. This system realized 65-megapixel videography of 12 wavebands within the visible wavelength range, with 25 nm spectral resolution. Besides, tunable filters such as Liquid Crystal Tunable Filters (LCTFs)¹⁵², Acousto-Optic Tunable Filters (AOTFs)¹⁵³, and Fabry-Perot Tunable Filters (FPTFs)¹⁵⁴ have also found applications in spectral modulation¹⁵⁵. For instance, Wang et al. employed LCTF to build a compressive spectral imaging camera¹⁵⁶, which used an accurate transmission model to represent LCTF's band-pass spectral filtering effects, with a randomly coded aperture placed behind it to modulate spectra in the spatial domain. With computational reconstruction, the camera overcame the intrinsic low spatial resolution limitation of traditional LCTF spectral systems (approximately 50×50 pixels¹⁵⁷) and achieved high-resolution multispectral imaging of 400×400 pixels in 500–710 nm with 10 nm spectral resolution. Although tunable filters are becoming increasingly smaller and more precise¹⁵⁸, their multiple acquisitions degrade imaging speed and limit their application. In most recent years, new tunable devices featuring micro-nano structures and dynamic filtering capabilities have emerged¹⁵⁹, which hold great potential for further applications in computational spectral imaging with end-to-end optimization.

Conclusion and discussion

This review lies in the joint encoding-decoding optimization framework for computational optics and discusses a series of modulation elements for physical twinning that implements optimized parameters regarding different modulation dimensions of amplitude,

phase, and spectrum. The concept of combining physical models with mathematical computation aligns with similar ideas in the fields of unrolling network¹⁶⁰ and physics-inspired network¹⁶¹, underscoring that interdisciplinary study is becoming a powerful technique to tackle various challenges. With the advancement of artificial intelligence, materials manufacturing, and physics modeling, end-to-end jointly optimized computational optics will assist in high-efficiency and reliable tasks in various fields including visual reality, remote sensing, biomedicine, etc.

In Table 1, we briefly outline the advantages and disadvantages of the representative modulation elements discussed in each modulation dimension, aiming to guide readers in selecting appropriate modulation devices for their specific task requirements. To better illustrate their features, we also list common performance parameters in Table 2. Furthermore, as shown in Table 3, we summarize their representative applications and include a column indicating whether each device can be adaptively designed. If so, these flexible devices can be applied for adaptive modulation tasks. In contrast, the modulation devices of fixed structures are less suitable for tasks requiring adaptive modulation. The table also shows that some devices, such as DMD and LED, are commonly limited to amplitude-dimensional imaging applications. In contrast, the devices such as SLM, DOE, and metasurfaces demonstrate powerful applications in both the phase and spectral dimensions.

Despite the above successful demonstrations, end-to-end jointly optimized computational optics and corresponding physical twinning still face several inherent challenges that necessitate further investigations.

- *A universal framework for end-to-end optimization*

The myriad imaging and sensing tasks exhibit various requirements in terms of computation time, network scale, decoding precision, and so on. Consequently, a delicate balance should be carefully considered between encoding efficiency and decoding quality, mandating distinct network structures for each task alongside comprehensive and repeated network learning stages. This leads to high computational complexity in the training stage, which presents an impediment to the widespread adoption of end-to-end joint optimization, particularly in scenarios with limited computational capabilities. To address this challenge, for instance, sampling-adaptive approaches^{59,162} focus on identifying common parameters while prioritizing adjustments solely to the sampling ratio, thereby minimizing alterations to a small subset of network parameters. We advocate for a deeper exploration of shared characteristics among diverse tasks and the development of a unified architecture that considers



Table 1 Comparison of different modulation elements in the dimensions of amplitude, phase, and spectrum

Modulation dimension	Modulation element	Advantages	Disadvantages
Amplitude	Deformable Micromirror Device (DMD) ^{19,22,44,46,50,51,54–60,167}	High spatial resolution, high-speed modulation, high reliability	Complex optical systems, high cost, limited grayscale modulation performance
	Liquid Crystal (LC) ^{63,105}	Low cost, easy to integrate	Slow response speed
	Light-Emitting Diode (LED) ^{74,75,168}	Good color performance, high-speed modulation	Limited spatial resolution, high-temperature sensitivity
Phase	Diffractive Optical Element (DOE) ^{92,169,170}	Low cost	Large scale
	Phase-only SLM ¹⁰³	High adaptability, flexible application	High cost, complex calibration
	Metasurface ^{30,171,172}	Highly customizable, high precision, ultra-thin design	High cost, sensitivity to environmental conditions
Spectrum	Filter array ^{136,137}	High spectral resolution, high flexibility	Finite bandwidth, low flexibility
	Diffractive Optical Element (DOE) ^{34,142,143}	Low cost, high throughput	Sensitivity to environmental conditions
	Spatial Light Modulator (SLM) ^{147,173,174}	High-speed modulation, high flexibility	Sensitivity to environmental conditions, high cost
	Metasurface ^{116,130,141,149–151}	Compact structure, efficient modulation	High cost, limited scalability

these common parameters. Such an approach would alleviate the computational burden associated with end-to-end joint optimization, thereby facilitating its broader implementation.

- *Information security*
Within the framework of end-to-end optimization, physical twinning represents the reverse implementation of the optimized encoding parameters. The modulation devices, illumination matrices, and other information related to this stage can be considered key parameters of the neural network layers, encompassing substantial inference information about the target¹⁶³. For certain applications such as medical diagnostics and remote sensing, information leakage can lead to significant and unexpected privacy issues. Therefore, in end-to-end optimized computational optics tasks, it is essential to raise significant concerns about protecting both the decoding network model and hardware information used in physical twinning.
- *Multi-dimensional joint optimization*
Although existing studies have demonstrated multi-dimensional joint imaging such as holographic video imaging¹⁶⁴ and hyperspectral video imaging¹⁶⁵, most of them neglect the high interrelationships among different dimensions. This encoding strategy not only leads to a substantial increase in the amount of

collected data but also results in notably high computational complexity and low efficiency in subsequent decoding operations. We advocate for further research efforts aimed at end-to-end joint optimization and global data compression across multiple dimensions. By utilizing a DNN structure for encoding simulations rather than relying on a single network layer or operation, discrete single-dimensional modulation elements could potentially be jointly optimized for multidimensional modulation, enhancing overall efficiency and performance.

- *Modulation element development*
Despite the various modulation elements introduced above, their performance is still not enough to afford the rising requirements of various applications. To better match the high-precision parameters optimized by neural networks, there is an increasing demand for delicate modulation elements with higher modulation resolution, bit depth, modulation efficiency, and stronger stability. In addition, as optical systems advance towards miniaturization and integration¹⁶⁶, micro-nano fabrication technologies and the development of intelligent multi-dimensional modulation elements will emerge as key focuses and challenges in the field of computational optics.

Table 2 Modulation performance of representative modulation elements

Modulation element	Model	Spatial resolution (pixels)	Modulation speed	Bit-depth / Phase modulation range	Efficiency	Spectral range
DMD	Texas DLP230NP	1920 × 1080	100 kHz (1 bit)	1–8 bit	–	400–700 nm
	Texas DLP9500UV	1920 × 1080	> 23 kHz (1 bit binary) and > 1700 Hz (8 bit gray)	1–8 bit	94% (fill factor)	363 to 420 nm
LC	Seacomp 128240C	240 × 128	64 Hz	1–24 bit	–	–
	KSXGA 088" M249 FLCOS	1280 × 1024	3.2 kHz for binary pattern and 85Hz for 24 bit video	1–24 bit	> 96% (fill factor)	430–700 nm
LED	LDM788AX (Xu et al. ⁷⁵)	32 × 32	1 kHz for binary pattern	8 bit	–	613–630 nm
	Salvador-Balaguer et al. ⁷⁶	32 × 32	40 kHz for binary pattern	8 bit	–	–
DOE	Beam shaper	1–10 μm	real time	0–2π	80–90%	–
	Grating	600–1800 lines/mm	real time	0–2π	30–60%	400–1000 nm
Filter array	Multispectral Filter Arrays ¹³⁵	512 × 512	real time	–	–	469–633 nm
	Broadband Encoding Stochastic (BEST) filters ¹³⁶	480 × 640	real time	–	16 filters, 40–60% fewer than required in previous studies	400–700 nm (1 nm)
Metasurface	Diffraction-based metasurface ¹⁵¹	800 × 800	real time	0–2π	~ 40%–80%	795–980 nm
	Metasurface projector ¹⁴⁹	~ 58,000	real time	0–2π	–	400–700 nm
Phase-only SLM	HOLOEYE PLUTO-2.1 SLM	1920 × 1080	60 Hz	0–2π	87% (fill factor)	420–850 nm
	Meadowlark Optics LCoS SLM	1920 × 1152	60 Hz	0–2π	93% (fill factor)	450–1550 nm



Table 3 Representative applications of different modulation elements in the end-to-end joint optimization framework

Modulation element	Representative applications	Dynamic adjustment
Deformable Micromirror Device (DMD)	Single-pixel imaging and sensing ^{19,22,44,45,50,52–54,56–58,59,61,62} , optical neural network ⁶⁰ , terahertz imaging ⁴⁶ , infrared imaging ^{48,49}	X
Liquid Crystal (LC)	Structured Illumination Microscopy (SIM) ⁶⁹ , lensless single-pixel imaging system ⁶³ , D ² NN ¹⁰³ , hyperspectral/ multispectral imaging ¹⁷⁵	X
Light-Emitting Diode (LED)	Single-pixel imaging ⁷⁵ , 3D holography ⁷⁶ , snapshot hyperspectral imaging ^{173,176}	X
Diffraction Optical Element (DOE)	Image deblurring ⁹² , 3D holography ^{93,177} , hyperspectral/multispectral imaging ^{139,142,143}	X
Metasurface	Holography ^{110–113} , D ² NN ^{30,114,115} , hyperspectral/multispectral imaging ^{149–151}	X
Filter array	Hyperspectral/multispectral imaging ^{134,135,137,178,179}	X

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Nos. 62131003, 62322502, 62088101), and the Guangdong Province Key Laboratory of Intelligent Detection in Complex Environment of Aerospace, Land and Sea (No. 2022KSYS016).

Author details

¹State Key Laboratory of CNS/ATM & MIT Key Laboratory of Complex-field Intelligent Sensing, Beijing Institute of Technology, Beijing & Zhuhai, China. ²Yangtze Delta Region Academy of Beijing Institute of Technology (Jiaxing), Jiaxing, China. ³School of Artificial Intelligence, Beijing Normal University, Beijing, China

Author contributions

L.B. and H.H. conceived the idea. X.Z. and L.B. drafted the manuscript. X.C. revised the phase modulation section. R.Y. revised the spectral modulation section. L.B., H.H., and J.Z. supervised the work and polished the manuscript.

Data availability

Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

Conflict of interest

L.B. and J.Z. hold patents on technologies related to this work (China patent numbers ZL 2019 1 0697301.7, ZL 2020 1 1295796.X, and ZL 2020 1 1157901.3) and submitted related patent applications.

Received: 3 April 2024 Revised: 13 January 2025 Accepted: 4 March 2025
Published online: 15 April 2025

References

- Barbastathis, G., Ozcan, A. & Situ, G. H. On the use of deep learning for computational imaging. *Optica* **6**, 921–943 (2019).
- Eils, R. & Athale, C. Computational imaging in cell biology. *J. Cell Biol.* **161**, 477–481 (2003).
- Ntziachristos, V. Going deeper than microscopy: the optical imaging frontier in biology. *Nat. Methods* **7**, 603–614 (2010).
- Hebden, J. C., Arridge, S. R. & Delpy, D. T. Optical imaging in medicine: I. Experimental techniques. *Phys. Med. Biol.* **42**, 825–840 (1997).
- Stender, A. S. et al. Single cell optical imaging and spectroscopy. *Chem. Rev.* **113**, 2469–2527 (2013).
- Waibel, A. et al. Phoneme recognition using time-delay neural networks. *IEEE Trans. Acoust. Speech, Signal Process.* **37**, 328–339 (1989).
- Elman, J. L. Finding structure in time. *Cogn. Sci.* **14**, 179–211 (1990).

- Minsky, M. L. Steps towards artificial intelligence. *Proc. Inst. Radio Eng.* **49**, 8–30 (1961).
- McCulloch, W. S. & Pitts, W. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys.* **5**, 115–133 (1943).
- Rosenblatt, F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychol. Rev.* **65**, 386–408 (1958).
- Rumelhart, D. E., Hinton, G. E. & Williams, R. J. Learning representations by back-propagating errors. *Nature* **323**, 533–536 (1986).
- Hunt, E. B. Artificial intelligence. (Amsterdam: Academic Press, 2014).
- Adler, A., Elad, M. & Zibulevsky, M. Compressed learning: A deep neural network approach. Print at <https://arxiv.org/abs/1610.09615> (2016).
- Sitzmann, V. et al. End-to-end optimization of optics and image processing for achromatic extended depth of field and super-resolution imaging. *ACM Trans. Graph.* **37**, 114 (2018).
- Zhou, H. Q. et al. Single-shot phase retrieval based on anisotropic metasurface. *Appl. Phys. Lett.* **120**, 161702 (2022).
- Shi, L., Li, B. C. & Matusik, W. End-to-end learning of 3D phase-only holograms for holographic display. *Light Sci. Appl.* **11**, 247 (2022).
- Xu, Y. B. et al. A compressive hyperspectral video imaging system using a single-pixel detector. *Nat. Commun.* **15**, 1456 (2024).
- Wetzstein, G. et al. Inference in artificial intelligence with deep optics and photonics. *Nature* **588**, 39–47 (2020).
- Bian, L. H. et al. Image-free multi-character recognition. *Opt. Lett.* **47**, 1343–1346 (2022).
- Zhang, Z. B. et al. Fast Fourier single-pixel imaging via binary illumination. *Sci. Rep.* **7**, 12029 (2017).
- Zhang, Z. B. et al. Hadamard single-pixel imaging versus Fourier single-pixel imaging. *Opt. Express* **25**, 19619–19639 (2017).
- Fu, H., Bian, L. H. & Zhang, J. Single-pixel sensing with optimal binarized modulation. *Opt. Lett.* **45**, 3111–3114 (2020).
- Thorlabs. SLM at https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=10378 (2019).
- Turtaev, S. et al. Comparison of nematic liquid-crystal and DMD based spatial light modulation in complex photonics. *Opt. Express* **25**, 29874–29884 (2017).
- Edgar, M. P., Gibson, G. M. & Padgett, M. J. Principles and prospects for single-pixel imaging. *Nat. Photonics* **13**, 13–20 (2019).
- Huang, H. et al. 25,000 fps computational ghost imaging with ultrafast structured illumination. *Electron. Mater.* **3**, 93–100 (2022).
- Shen, Y. C. et al. Deep learning with coherent nanophotonic circuits. *Nat. Photonics* **11**, 441–446 (2017).
- Fukui, T., Nakano, Y. & Tanemura, T. On ghost imaging using multimode fiber and integrated optical phased array. Proceedings of the 24th OptoElectronics and Communications Conference (OECC) and 2019 International Conference on Photonics in Switching and Computing (PSC) 1–3 (IEEE, 2019).
- Bogaerts, W. et al. Programmable photonic circuits. *Nature* **586**, 207–216 (2020).
- Luo, X. H. et al. Metasurface-enabled on-chip multiplexed diffractive neural networks in the visible. *Light Sci. Appl.* **11**, 158 (2022).



31. Bernstein, L. et al. Single-shot optical neural network. *Sci. Adv.* **9**, eadg7904 (2023).
32. Hagen, N. A. & Kudenov, M. W. Review of snapshot spectral imaging technologies. *Opt. Eng.* **52**, 090901 (2013).
33. Lin, X. et al. Spatial-spectral encoded compressive hyperspectral imaging. *ACM Trans. Graph.* **33**, 233 (2014).
34. Baek, S. H. et al. Single-shot hyperspectral-depth imaging with learned diffractive optics. Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. 2651–2660 (IEEE, 2021).
35. Hua, X. et al. Ultra-compact snapshot spectral light-field imaging. *Nat. Commun.* **13**, 2732 (2022).
36. Mias, S. & Camon, H. A review of active optical devices: I. Amplitude modulation. *J. Micromech. Microeng.* **18**, 083001 (2008).
37. Brady, D. J. et al. Multiscale gigapixel photography. *Nature* **486**, 386–389 (2012).
38. Akbulut, D. et al. Focusing light through random photonic media by binary amplitude modulation. *Opt. Express* **19**, 4017–4029 (2011).
39. Tao, X. D. et al. High-speed scanning interferometric focusing by fast measurement of binary transmission matrix for channel demixing. *Opt. Express* **23**, 14168–14187 (2015).
40. Rabut, C. et al. Ultrafast amplitude modulation for molecular and hemodynamic ultrasound imaging. *Appl. Phys. Lett.* **118**, 244102 (2021).
41. Widrow, B. A study of rough amplitude quantization by means of Nyquist sampling theory. *IRE Trans. Circuit Theory* **3**, 266–276 (1956).
42. Zisselman, E., Adler, A. & Elad, M. Compressed learning for image classification: A deep neural network approach. *Handb. Numer. Anal.* **19**, 3–17 (2018).
43. LeCun, Y. et al. Gradient-based learning applied to document recognition. *Proc. IEEE* **86**, 2278–2324 (1998).
44. Higham, C. F. et al. Deep learning for real-time single-pixel video. *Sci. Rep.* **8**, 2369 (2018).
45. Yang, Z. H. et al. Image-free real-time target tracking by single-pixel detection. *Opt. Express* **30**, 864–873 (2022).
46. Deng, Y. et al. High-efficiency terahertz single-pixel imaging based on a physics-enhanced network. *Opt. Express* **31**, 10273–10286 (2023).
47. Hornbeck, L. J. 128 × 128 deformable mirror device. *IEEE Trans. Electron Devices* **30**, 539–545 (1983).
48. Gattinger, P. et al. Spectral-coding-based compressive single-pixel NIR spectroscopy in the sub-millisecond regime. *Sensors* **21**, 5563 (2021).
49. Ebner, A. et al. Diffraction-limited hyperspectral mid-infrared single-pixel microscopy. *Sci. Rep.* **13**, 281 (2023).
50. Johnson, S. D. et al. Single-pixel RIDAR with deep learning optimised sampling. In *Proc. of the 2020 Conference on Lasers and Electro-Optics*. 1–2 (IEEE, 2020).
51. Hoshi, I. et al. Single-pixel imaging using a recurrent neural network combined with convolutional layers. *Opt. Express* **28**, 34069–34078 (2020).
52. Jiang, W. et al. Imaging high-speed moving targets with a single-pixel detector. *Opt. Express* **28**, 7889–7897 (2020).
53. Xu, M. et al. Deep compressed super-resolution imaging with DMD alignment error correction. *Photonics* **10**, 581 (2023).
54. Liu, H. Y. et al. Masked autoencoder for highly compressed single-pixel imaging. *Opt. Lett.* **48**, 4392–4395 (2023).
55. Zhan, X. R. et al. Scattering-induced entropy boost for highly-compressed optical sensing and encryption. Print at <https://arxiv.org/abs/2301.06084> (2022).
56. Liu, H. Y., Bian, L. H. & Zhang, J. Image-free single-pixel segmentation. *Opt. Laser Technol.* **157**, 108600 (2023).
57. Peng, L. T. et al. Image-free single-pixel object detection. *Opt. Lett.* **48**, 2527–2530 (2023).
58. Zhan, X. R. et al. Ultrahigh-security single-pixel semantic encryption. *Opt. Lett.* **47**, 6169–6172 (2022).
59. Zhan, X. R. et al. Weighted sampling-adaptive single-pixel sensing. *Opt. Lett.* **47**, 2838–2841 (2022).
60. Jiao, S. M. et al. Optical machine learning with incoherent light and a single-pixel detector. *Opt. Lett.* **44**, 5186–5189 (2019).
61. Zhan, X. et al. Global-optimal semi-supervised learning for single-pixel image-free sensing. *Opt. Lett.* **49**, 682–685 (2024).
62. Wan, S. B., Chang, X. Y. & Bian, L. H. Image-free classification via few-shot learning. Proceedings of SPIE 12767, Optoelectronic Imaging and Multimedia Technology X. 229–238 (SPIE, 2023).
63. Zhang, Z. B. et al. Lensless single-pixel imaging by using LCD: application to small-size and multi-functional scanner. *Opt. Express* **27**, 3731–3745 (2019).
64. Bleha, Jr. W. P. & Lei, L. A. Advances in liquid crystal on silicon (LCoS) spatial light modulator technology. Proceedings of SPIE 8736, Display Technologies and Applications for Defense, Security, and Avionics VII. 47–54 (SPIE, 2013).
65. Janssen, P. et al. Design aspects of a scrolling color LCoS display. *Displays* **23**, 99–108 (2002).
66. Kner, P. et al. Super-resolution video microscopy of live cells by structured illumination. *Nat. Methods* **6**, 339–342 (2009).
67. Shao, L. et al. Super-resolution 3D microscopy of live whole cells using structured illumination. *Nat. Methods* **8**, 1044–1046 (2011).
68. Lu-Walther, H. W. et al. fastSIM: a practical implementation of fast structured illumination microscopy. *Methods Appl. Fluoresc.* **3**, 014001 (2015).
69. Hannebelle, M. T. M. et al. Open-source microscope add-on for structured illumination microscopy. *Nat. Commun.* **15**, 1550 (2024).
70. Demmerle, J. et al. Strategic and practical guidelines for successful structured illumination microscopy. *Nat. Protoc.* **12**, 988–1010 (2017).
71. Andreev, A. L. et al. Novel FLC-materials open new possibilities for FLCoS based microdisplays and video projectors. *Phys. Procedia* **73**, 87–94 (2015).
72. Yang, D. K. & Wu, S. T. Fundamentals of Liquid Crystal Devices. 2nd edn. (Chichester: John Wiley & Sons, 2015).
73. Pershan, P. S. Structure of Liquid Crystal Phases (World Scientific, 1988).
74. Mizuno, T. & Iwata, T. Hadamard-transform fluorescence-lifetime imaging. *Opt. Express* **24**, 8202–8213 (2016).
75. Xu, Z. H. et al. 1000 fps computational ghost imaging using LED-based structured illumination. *Opt. Express* **26**, 2427–2434 (2018).
76. Salvador-Balaguer, E. et al. Low-cost single-pixel 3D imaging by using an LED array. *Opt. Express* **26**, 15623–15631 (2018).
77. Hahamovich, E. et al. Single pixel imaging at megahertz switching rates via cyclic Hadamard masks. *Nat. Commun.* **12**, 4516 (2021).
78. He, Y. H. et al. Single-pixel imaging with neutrons. *Sci. Bull.* **66**, 133–138 (2021).
79. Goodman, J. W. Introduction to Fourier Optics. 3rd edn. (Englewood: Roberts and Company publishers, 2005).
80. Park, Y., Depeursinge, C. & Popescu, G. Quantitative phase imaging in biomedicine. *Nat. Photonics* **12**, 578–589 (2018).
81. Chang, X. Y., Bian, L. H. & Zhang, J. Large-scale phase retrieval. *eLight* **1**, 4 (2021).
82. Yang, W. J. & Yuste, R. In vivo imaging of neural activity. *Nat. Methods* **14**, 349–359 (2017).
83. Booth, M. J. Adaptive optics in microscopy. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **365**, 2829–2843 (2007).
84. Yu, H. et al. Recent advances in wavefront shaping techniques for biomedical applications. *Curr. Appl. Phys.* **15**, 632–641 (2015).
85. Lin, X. et al. All-optical machine learning using diffractive deep neural networks. *Science* **361**, 1004–1008 (2018).
86. Sui, X. et al. A review of optical neural networks. *IEEE Access* **8**, 70773–70783 (2020).
87. Piechulla, P. M. et al. Antireflective Huygens' metasurface with correlated disorder made from high-index disks implemented into silicon heterojunction solar cells. *ACS Photonics* **8**, 3476–3485 (2021).
88. Chang, N. Y. & Kuo, C. J. Algorithm based on rigorous coupled-wave analysis for diffractive optical element design. *J. Opt. Soc. Am. A* **18**, 2491–2501 (2001).
89. Prather, D. W. & Shi, S. Formulation and application of the finite-difference time-domain method for the analysis of axially symmetric diffractive optical elements. *J. Opt. Soc. Am.* **16**, 1131–1142 (1999).
90. Taghizadeh, M. R. et al. Design and fabrication of diffractive optical elements. *Microelectron. Eng.* **34**, 219–242 (1997).
91. Khonina, S. N., Kazanskiy, N. L. & Butt, M. A. Exploring diffractive optical elements and their potential in free space optics and imaging—a comprehensive review. *Laser Photonics Rev.* **18**, 2400377 (2024).
92. Li, Y. Q., Fu, Q. & Heidrich, W. Extended depth-of-field projector using learned diffractive optics. Proceedings of 2023 IEEE Conference Virtual Reality and 3D User Interfaces. 449–459 (IEEE, 2023).
93. Shi, L. et al. Towards real-time photorealistic 3D holography with deep neural networks. *Nature* **591**, 234–239 (2021).
94. Bawart, M., Bernet, S. & Ritsch-Marte, M. Programmable freeform optical elements. *Opt. Express* **25**, 4898–4906 (2017).



95. Swanson, G. J. Binary optics technology: theoretical limits on the diffraction efficiency of multilevel diffractive optical elements. (Massachusetts Institute of Technology, 1991).
96. Bauchert, K. A., Serati, S. A. & Furman, A. Advances in liquid crystal spatial light modulators. *Proceedings of SPIE 4734, Optical Pattern Recognition XIII*. 35–43 (SPIE, 2002).
97. Takaki, Y. & Fujii, K. Viewing-zone scanning holographic display using a mems spatial light modulator. *Opt. Express* **22**, 24713–24721 (2014).
98. Igasaki, Y. et al. High efficiency electrically-addressable phase-only spatial light modulator. *Opt. Rev.* **6**, 339–344 (1999).
99. Hashimoto, N., Morokawa, S. & Kitamura, K. Real-time holography using the high-resolution LCTV-SLM. *Proceedings of SPIE 1461, Practical Holography V*. 291–302 (SPIE, 1991).
100. Salter, P. S. & Booth, M. J. Adaptive optics in laser processing. *Light Sci. Appl.* **8**, 110 (2019).
101. Li, S. S. et al. High-quality near-field beam achieved in a high-power laser based on SLM adaptive beam-shaping system. *Opt. Express* **23**, 681–689 (2015).
102. Chen, R. Y. et al. Physics-aware machine learning and adversarial attack in complex-valued reconfigurable diffractive all-optical neural network. *Laser Photonics Rev.* **16**, 2200348 (2022).
103. Wang, H. et al. Intelligent optoelectronic processor for orbital angular momentum spectrum measurement. *Photonix* **4**, 9 (2023).
104. Pinilla, S. et al. Hybrid diffractive optics design via hardware-in-the-loop methodology for achromatic extended-depth-of-field imaging. *Opt. Express* **30**, 32633–32649 (2022).
105. Lazarev, G. et al. LCOS spatial light modulators: trends and applications. in *Optical Imaging and Metrology: Advanced Technologies* (eds Osten, W. & Reingand, N.) 1–29 (Wiley-VCH, 2012).
106. Bowman, R. et al. Optimisation of a low cost SLM for diffraction efficiency and ghost order suppression. *Eur. Phys. J. Spec. Top.* **199**, 149–158 (2011).
107. Gao, Y. H., Li, R. J. & Cao, L. C. Self-referenced multiple-beam interferometric method for robust phase calibration of spatial light modulator. *Opt. Express* **27**, 34463–34471 (2019).
108. Chen, H. T., Taylor, A. J. & Yu, N. F. A review of metasurfaces: physics and applications. *Rep. Prog. Phys.* **79**, 076401 (2016).
109. Hu, J. et al. A review on metasurface: from principle to smart metadevices. *Front. Phys.* **8**, 586087 (2021).
110. Zheng, G. X. et al. Metasurface holograms reaching 80% efficiency. *Nat. Nanotechnol.* **10**, 308–312 (2015).
111. Ye, W. M. et al. Spin and wavelength multiplexed nonlinear metasurface holography. *Nat. Commun.* **7**, 11930 (2016).
112. Li, L. L. et al. Electromagnetic reprogrammable coding-metasurface holograms. *Nat. Commun.* **8**, 197 (2017).
113. Huang, L. L., Zhang, S. & Zentgraf, T. Metasurface holography: from fundamentals to applications. *Nanophotonics* **7**, 1169–1190 (2018).
114. Tseng, E. et al. Neural nano-optics for high-quality thin lens imaging. *Nat. Commun.* **12**, 6493 (2021).
115. Liu, C. et al. A programmable diffractive deep neural network based on a digital-coding metasurface array. *Nat. Electron.* **5**, 113–122 (2022).
116. Su, V. C. et al. Advances in optical metasurfaces: fabrication and applications [Invited]. *Opt. Express* **26**, 13148–13182 (2018).
117. He, Q. et al. High-efficiency metasurfaces: principles, realizations, and applications. *Adv. Opt. Mater.* **6**, 1800415 (2018).
118. Moitra, P. et al. Large-scale all-dielectric metamaterial perfect reflectors. *ACS Photonics* **2**, 692–698 (2015).
119. Ledwig, P. & Robles, F. E. Epi-mode tomographic quantitative phase imaging in thick scattering samples. *Biomed. Opt.* **10**, 3605–3621 (2019).
120. Georgieva, A., Belashov, A. V. & Petrov, N. V. Optimization of DMD-based independent amplitude and phase modulation by analysis of target complex wavefront. *Sci. Rep.* **12**, 7754 (2022).
121. Goorden, S. A., Bertolotti, J. & Mosk, A. P. Superpixel-based spatial amplitude and phase modulation using a digital micromirror device. *Opt. Express* **22**, 17999–18009 (2014).
122. Hou, H. Y. et al. Complex-amplitude single-pixel imaging using coherent structured illumination. *Opt. Express* **29**, 41827–41841 (2021).
123. Newton, I. A letter of Mr. Isaac Newton, professor of the mathematics in the university of Cambridge; containing his new theory about light and colors: sent by the author to the publisher from Cambridge, Febr. 6. 1671²¹ in order to be communicated to the R. society. *Philos. Trans.* **6**, 3075–3087 (1672).
124. Vojtišek, P., Květoň, M. & Richter, I. Effective spectral dispersion of refractive index modulation. *J. Opt.* **19**, 045603 (2017).
125. Huang, L. Q. et al. Spectral imaging with deep learning. *Light Sci. Appl.* **11**, 61 (2022).
126. Garini, Y., Young, I. T. & McNamara, G. Spectral imaging: Principles and applications. *Cytom. Part A* **69A**, 735–747 (2006).
127. Arguello, H. et al. Deep optical coding design in computational imaging: a data-driven framework. *IEEE Signal Process. Mag.* **40**, 75–88 (2023).
128. Manolakis, D. & Shaw, G. Detection algorithms for hyperspectral imaging applications. *IEEE Signal Process. Mag.* **19**, 29–43 (2002).
129. Khan, M. J. et al. Modern trends in hyperspectral image analysis: a review. *IEEE Access* **6**, 14118–14129 (2018).
130. Li, T. et al. Revolutionary meta-imaging: from superlens to metalens. *Photonics Insights* **2**, R01 (2023).
131. Anzagira, L. & Fossum, E. R. Color filter array patterns for small-pixel image sensors with substantial cross talk. *J. Opt. Soc. Am. A* **32**, 28–34 (2015).
132. Chou Chau, Y. F. et al. A multichannel color filter with the functions of optical sensor and switch. *Sci. Rep.* **11**, 22910 (2021).
133. Wang, S. W. et al. Concept of a high-resolution miniature spectrometer using an integrated filter array. *Opt. Lett.* **32**, 632–634 (2007).
134. Mihoubi, S. et al. Multispectral demosaicing using pseudo-panchromatic image. *IEEE Trans. Comput. Imaging* **3**, 982–995 (2017).
135. Zhang, T., Liang, Z. Y. & Fu, Y. Joint spatial-spectral pattern optimization and hyperspectral image reconstruction. *IEEE J. Sel. Top. Signal Process.* **16**, 636–648 (2022).
136. Zhang, W. Y. et al. Deeply learned broadband encoding stochastic hyperspectral imaging. *Light Sci. Appl.* **10**, 108 (2021).
137. Song, H. Y. et al. Deep-learned broadband encoding stochastic filters for computational spectroscopic instruments. *Adv. Theory Simul.* **4**, 2000299 (2021).
138. Wu, R. J. et al. Optimized multi-spectral filter arrays for spectral reconstruction. *Sensors* **19**, 2905 (2019).
139. Arguello, H. et al. Shift-variant color-coded diffractive spectral imaging system. *Optica* **8**, 1424–1434 (2021).
140. Chen, C. C., Gu, H. G. & Liu, S. Y. Ultra-simplified diffraction-based computational spectrometer. *Light Sci. Appl.* **13**, 9 (2024).
141. Engelberg, J. & Levy, U. The advantages of metalenses over diffractive lenses. *Nat. Commun.* **11**, 1991 (2020).
142. Lv, T. et al. Aperture diffraction for compact snapshot spectral imaging. *Proceedings of the 2023 IEEE/CVF International Conference on Computer Vision*. 10574–10584 (IEEE, 2023).
143. Mengu, D. et al. Snapshot multispectral imaging using a diffractive optical network. *Light Sci. Appl.* **12**, 86 (2023).
144. Shi, Z. et al. Compact self-adaptive coding for spectral compressive sensing. *Proceeding of 2023 IEEE International Conference on Computational Photography*. 1–12 (IEEE, 2023).
145. Li, Z. W. et al. Efficient single-pixel multispectral imaging via non-mechanical spatio-spectral modulation. *Sci. Rep.* **7**, 41435 (2017).
146. Arce, G. R. et al. Compressive coded aperture spectral imaging: an introduction. *IEEE Signal Process. Mag.* **31**, 105–115 (2014).
147. Wang, L. Z. et al. HyperReconNet: joint coded aperture optimization and image reconstruction for compressive hyperspectral imaging. *IEEE Trans. Image Process.* **28**, 2257–2270 (2019).
148. Zhang, W. H. et al. Handheld snapshot multi-spectral camera at tens-of-megapixel resolution. *Nat. Commun.* **14**, 5043 (2023).
149. McClung, A. et al. Snapshot spectral imaging with parallel metasystems. *Sci. Adv.* **6**, eabc7646 (2020).
150. Makarenko, M. et al. Real-time hyperspectral imaging in hardware via trained metasurface encoders. In *Proc. of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 12692–12702 (IEEE, 2022).
151. Zhang, Q. B. et al. End-to-end joint optimization of metasurface and image processing for compact snapshot hyperspectral imaging. *Opt. Commun.* **530**, 129154 (2023).
152. Hardeberg, J. Y., Schmitt, F. J. M. & Brettel, H. Multispectral color image capture using a liquid crystal tunable filter. *Opt. Eng.* **41**, 2532–2548 (2002).
153. Chang, I. C. Acousto-optic tunable filters. *Opt. Eng.* **20**, 206824 (1981).
154. Neumann, N. et al. Tunable infrared detector with integrated micro-machined Fabry-Perot filter. *J. Micro/Nanolithogr. MEMS, MOEMS* **7**, 021004 (2008).



155. López-Álvarez, M. A., Hernández-Andrés, J. & Romero, J. Developing an optimum computer-designed multispectral system comprising a mono-chrome ccd camera and a liquid-crystal tunable filter. *Appl. Opt.* **47**, 4381–4390 (2008).
156. Wang, X. et al. Compressive spectral imaging system based on liquid crystal tunable filter. *Opt. Express* **26**, 25226–25243 (2018).
157. Gebhart, S. C., Thompson, R. C. & Mahadevan-Jansen, A. Liquid-crystal tunable filter spectral imaging for brain tumor demarcation. *Appl. Opt.* **46**, 1896–1910 (2007).
158. Xie, Z. W. et al. Liquid-crystal tunable color filters based on aluminum metasurfaces. *Opt. Express* **25**, 30764–30770 (2017).
159. He, Q. et al. Dynamically tunable transmissive color filters using ultra-thin phase change materials. *Opt. Express* **28**, 39841–39849 (2020).
160. Monga, V., Li, Y. L. & Eldar, Y. C. Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing. *IEEE Signal Process. Mag.* **38**, 18–44 (2021).
161. Zhang, J. et al. Physics-inspired compressive sensing: Beyond deep unrolling. *IEEE Signal Process. Mag.* **40**, 58–72 (2023).
162. Phillips, D. B. et al. Adaptive foveated single-pixel imaging with dynamic supersampling. *Sci. Adv.* **3**, e1601782 (2017).
163. Fredrikson, M., Jha, S. & Ristenpart, T. Model inversion attacks that exploit confidence information and basic countermeasures. In *Proc. of the 22nd ACM SIGSAC Conference on Computer and Communications Security*. 1322–1333 (ACM, 2015).
164. Shi, Z. J. et al. Super-resolution orbital angular momentum holography. *Nat. Commun.* **14**, 1869 (2023).
165. Martel, J. N. P. et al. Neural sensors: Learning pixel exposures for HDR imaging and video compressive sensing with programmable sensors. *IEEE Trans. Pattern Anal. Mach. Intell.* **42**, 1642–1653 (2020).
166. Yang, Z. Y. et al. Miniaturization of optical spectrometers. *Science* **371**, eabe0722 (2021).
167. Ma, M. C. et al. Physics-informed and simulation-driven optimization for binary Fourier single-pixel imaging. *Opt. Lett.* **49**, 218–221 (2024).
168. Horstmeyer, R. et al. Convolutional neural networks that teach microscopes how to image. Print at <https://arxiv.org/abs/1709.07223> (2017).
169. Kellman, M. R. et al. Physics-based learned design: optimized coded-illumination for quantitative phase imaging. *IEEE Trans. Comput. Imaging* **5**, 344–353 (2019).
170. Muthumbi, A. et al. Learned sensing: jointly optimized microscope hardware for accurate image classification. *Biomed. Opt. Express* **10**, 6351–6369 (2019).
171. Ghorbani, F. et al. Deep neural network-based automatic metasurface design with a wide frequency range. *Sci. Rep.* **11**, 7102 (2021).
172. Banerji, S. et al. Imaging with flat optics: metalenses or diffractive lenses? *Optica* **6**, 805–810 (2019).
173. Fu, Y. et al. Coded hyperspectral image reconstruction using deep external and internal learning. *IEEE Trans. Pattern Anal. Mach. Intell.* **44**, 3404–3420 (2022).
174. Meng, Z. Y., Ma, J. W. & Yuan, X. End-to-end low cost compressive spectral imaging with spatial-spectral self-attention. In *Proc. of the 16th European Conference on Computer Vision*. 187–204 (Springer, 2020).
175. Lin, X. et al. Dual-coded compressive hyperspectral imaging. *Opt. Lett.* **39**, 2044–2047 (2014).
176. Zhang, S. P. et al. Learning tensor low-rank prior for hyperspectral image reconstruction. In *Proc. of 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 12006–12015 (IEEE, 2021).
177. Işıl, Ç. et al. All-optical image denoising using a diffractive visual processor. *Light Sci. Appl.* **13**, 43 (2024).
178. Nie, S. J. et al. Deeply learned filter response functions for hyperspectral reconstruction. In *Proc. of the 2018 IEEE Conference on Computer Vision and Pattern Recognition*. (IEEE, 2018).
179. Gharbi, M. et al. Deep joint demosaicking and denoising. *ACM Trans. Graph. (TOG)* **35**, 191 (2016).
180. Wikipedia. Computational imaging (2019). at <https://en.wikipedia.org/wiki/Computationalimaging>.
181. Sinha, A. et al. Lensless computational imaging through deep learning. *Optica* **4**, 1117–1125 (2017).
182. Bai, B. J. et al. Pyramid diffractive optical networks for unidirectional image magnification and demagnification. *Light Sci. Appl.* **13**, 178 (2024).
183. Huang, Y. G. et al. Liquid-crystal-on-silicon for augmented reality displays. *Appl. Sci.* **8**, 2366 (2018).
184. Sun, M. J. et al. Lensless imaging via LED array based computational ghost imaging. *Opt. Laser Technol.* **180**, 111401 (2025).
185. Allen, J. Application of patterned illumination using a DMD for optogenetic control of signaling. *Nat. Methods* **14**, 1114 (2017).
186. Li, R. J. & Cao, L. C. Progress in phase calibration for liquid crystal spatial light modulators. *Appl. Sci.* **9**, 2012 (2019).
187. Tian, P. F. et al. Fabrication, characterization and applications of flexible vertical InGaN micro-light emitting diode arrays. *Opt. Express* **24**, 699–707 (2016).
188. Shen, C. Y. et al. All-optical phase conjugation using diffractive wavefront processing. *Nat. Commun.* **15**, 4989 (2024).
189. Hyman, R. M. et al. Polarization-independent phase modulation using a blue-phase liquid crystal over silicon device. *Appl. Opt.* **15**, 6925–6929 (2024).
190. Li, J. et al. Mechanisms of 2π phase control in dielectric metasurface and transmission enhancement effect. *Opt. Express* **27**, 23186–23196 (2019).
191. Kittle, D. et al. Multiframe image estimation for coded aperture snapshot spectral imagers. *Appl. Opt.* **49**, 6824–6833 (2010).
192. Xu, N. et al. Snapshot hyperspectral imaging based on equalization designed DOE. *Opt. Express* **31**, 20489–20504 (2023).

