

Review

Towards Next-Generation Smart Seed Phenomics: A Review and Roadmap for Metasurface-Based Hyperspectral Imaging and a Light-Field Platform for 3D Reconstruction

Jingrui Yang ^{1,2}, Qinglei Zhao ¹, Shuai Liu ^{1,2}, Jing Guo ^{1,2}, Fengwei Guan ¹ , Shuxin Wang ¹ , Qinglong Hu ¹, Qiang Liu ¹, Qi Song ^{3,*} , Mingdong Zhu ^{4,5}  and Chao Li ^{6,*} 

¹ Changchun Institute of Optics, Fine Mechanics and Physics, CAS, Changchun 130033, China; 18043063676@163.com (J.Y.); coldsun@ciomp.ac.cn (Q.Z.); guo_jing_1234@163.com (J.G.); buaaguan@sina.com (F.G.); wangsx@ciomp.ac.cn (S.W.); huql@ciomp.ac.cn (Q.H.); liuqiang@ciomp.ac.cn (Q.L.)

² University of the Chinese Academy of Sciences, Beijing 100039, China

³ Soleilware Photonics LLC, San Diego, CA 92131, USA

⁴ State Key Laboratory of Hybrid Rice, Hunan Academy of Agricultural Sciences, Changsha 410125, China; uhz_uhz@hotmail.com

⁵ Yuelushan Laboratory, Changsha 410128, China

⁶ Qualcomm, San Diego, CA 92131, USA

* Correspondence: songqi@sibet.ac.cn (Q.S.); snappylichao@gmail.com (C.L.)

Abstract

Seed phenomics is a critical research field for understanding seed germination mechanisms. Metasurfaces, composed of subwavelength nanostructures, offer a promising pathway to achieve both dispersion control and imaging functionalities within an ultra-compact form factor. Recent advances in micro–nano-optics and computational imaging have opened new avenues for high-dimensional, multimodal imaging. However, conventional hyperspectral and light-field systems still face limitations in compactness, depth resolution, and spectral–spatial integration. This review summarizes recent progress in metalens and metasurface lens array-based light-field systems for hyperspectral imaging and 3D reconstruction, with a focus on the underlying principles, design strategies, and reconstruction algorithms that enable single-shot 3D hyperspectral acquisition. We further present a forward-looking roadmap toward the realization of a revolutionized imaging paradigm: a metasurface-based light-field platform that fully integrates 3D and hyperspectral imaging capabilities. In particular, we examine how dispersive metasurfaces serve as core optical elements for precise dispersion control in hyperspectral imaging systems, while metalens arrays enable accurate modulation of spatial–angular distributions in light-field configurations. We systematically review both 3D and spectral reconstruction algorithms, highlighting their roles in decoding complex optical encodings. The application of these integrated systems in seed phenotyping is emphasized, demonstrating their capability to capture 3D spatial–spectral distributions in a single exposure. This approach facilitates high-throughput analysis of morphological traits, germination potential, and internal biochemical composition, offering a comprehensive solution for advanced seed characterization. Finally, we outline a practical roadmap for implementing a metasurface-based light-field platform that integrates hyperspectral imaging and computational 3D reconstruction. This review offers a comprehensive overview of the state of the art in compact 3D light-field systems and multimodal hyperspectral imaging platforms, while providing forward-looking insights aimed at advancing smart seed phenotyping, precision agriculture, and next-generation optical imaging technologies.



Received: 25 November 2025

Revised: 29 December 2025

Accepted: 5 January 2026

Published: 8 January 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

Keywords: metasurfaces; 3D reconstruction; computational imaging; light-field imaging; hyperspectral imaging; seed phenomics; AI-driven phenotyping

1. Introduction

Rice and maize are among the world’s most important staple crops, collectively serving as primary food sources for nearly half of the global population. They are not only the world’s largest cereal by production volume, but also key raw materials spanning food, feed, and diverse industrial applications [1,2]. Their deep-processing sectors encompass multiple downstream industries, including pharmaceuticals and chemicals, making them increasingly significant drivers of global economic growth. Consequently, advancing our understanding of seed germination mechanisms and physiological attributes carries substantial strategic importance.

Seed phenomics seeks to elucidate the complex interplay between seed morphological, physiological, biochemical, and molecular characteristics and their underlying genetic determinants as well as environmental influences. Although recent advances in molecular biology, genomics, genetics, and bioinformatics have greatly accelerated progress in crop research, research on seed phenotyping has lagged behind. The remarkable diversity of crop species and the sheer scale of data required overwhelm traditional manual methods, which lack scalability, efficiency, and objectivity. As a result, there is a pressing need for phenotyping technologies that are automated, high-throughput, environmentally robust, platform-independent, and free from human bias.

Figure 1a presents high-definition three-dimensional visualizations of the surface morphology of rice seeds across developmental stages. Figure 1b,c show maize seed placement and hyperspectral imaging, which has been used to study seed vigor widely. Moreover, integrating 3D phenotypic information with hyperspectral data enables the construction of a quantitative “bridge” linking seed phenotypes to genomic datasets, thereby accelerating genetic improvement and breeding efforts while enhancing seed stress resilience. Spectral signatures provide a precise, non-destructive optical modality for probing the internal structural and biochemical composition of seeds. Using hyperspectral imaging, the spectral responses of rice seeds under varying temperature conditions can be characterized, offering a means to evaluate developmental status under fluctuating diurnal temperature regimes. Furthermore, the combination of multispectral imaging and machine learning facilitates efficient seed trait identification and developmental tracking based on infrared spectral features [3–5].

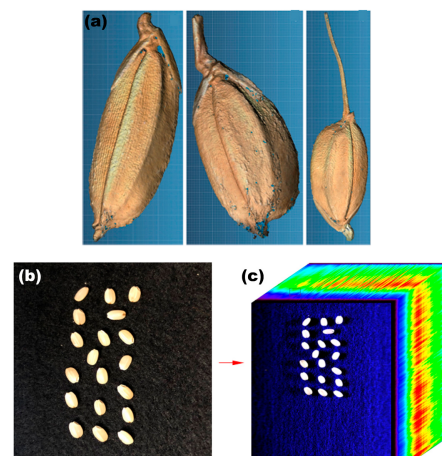


Figure 1. (a) High-resolution 3D surface visualization of rice seeds during development; (b) seed sample placement; (c) hypercube generation [6].

Therefore, three-dimensional seed phenotyping and spectral imaging technologies both play a pivotal role in internal structural analysis, germination monitoring, phenomic characterization, quality assessment, and dynamic developmental surveillance [7–10]. Moreover, the integration of hyperspectral and 3D imaging, through either spectrum-featured data and spatial information or the fusion of multidimensional “spectral–morphological–functional” data, provides unprecedented depth and breadth of information for crop breeding research. Therefore, developing specialized 3D hyperspectral imaging technologies for seeds represents a critical frontier with exceptional promise for advancing seed phenomics.

1.1. Challenges and Technological Gaps in High-Throughput Seed Phenomics

With the rapid advancement of sensor technologies, machine vision, and automation, non-destructive and high-throughput imaging techniques, spectral analysis, and image-processing methods have become increasingly mature. Coupled with powerful bioinformatics and data analytics capabilities, phenomics has emerged as a key frontier for enhancing breeding efficiency.

Nevertheless, seed phenotyping faces significant challenges. Many phenotypic traits exhibit subtle variations that are easily confounded by measurement tools, environmental fluctuations, or human subjectivity, making fine-scale features difficult to capture reliably. In addition, the measurable window for certain phenotypic traits is narrow. Achieving precise, standardized, and quantifiable phenotypic acquisition has therefore become a central challenge in agricultural production, crop breeding, and seed quality assessment [11].

To address these needs, numerous high-throughput phenotyping platforms have been developed worldwide. For example, the PhenoAIxpert HT (LemnaTec, Aachen, Germany) integrates a laser-based 3D imaging module capable of capturing plant surface point clouds non-destructively [12]; the TraitMill™ system (CropDesign, Ghent, Belgium) provides dynamic, continuous, large-scale data on genotype–phenotype–environment interactions [13]; the Crop3D platform developed by the Institute of Botany, Chinese Academy of Sciences, integrates LiDAR, thermal imaging, and hyperspectral sensing for precise 3D phenotyping [14]; researchers at Huazhong Agricultural University have developed a CT–RGB dual-modality system enabling *in vivo* monitoring of rice tillering with a resolution of 30 μm [15]; and Li et al. developed a point-cloud-based method using laser scanning to calculate seven morphological and geometric traits of rice seeds, including length, width, thickness, elongation, surface area, volume, and shape factors [16].

Despite these advances, current technologies still exhibit several critical limitations. First, conventional 2D imaging captures only planar morphological parameters and lacks true 3D shape, depth, and fine microstructural texture information, resulting in insufficient phenotypic dimensionality. Second, multi-camera array systems for 3D reconstruction require complex hardware synchronization and calibration; parameter inconsistencies across cameras often introduce registration errors and increase the difficulty of algorithm optimization. Third, single-modality imaging cannot fully characterize the relationships between seed physiological properties and developmental potential, thereby constraining the generalizability and predictive power of phenotypic modeling.

1.2. Light-Field and Metasurface Imaging: Emerging Technologies for 3D and Spectral Seed Phenotyping

To overcome the limitations of existing phenotyping systems, recent research has introduced light-field imaging and multimodal data fusion as a promising new paradigm for seed phenomics. This paradigm offers three major innovations. First, microlens-array light-field imaging enables multi-view, single-camera acquisition. By capturing both spatial and angular information of light rays, the system preserves standard 2D appearance

features while extracting 3D surface descriptors, such as curvature, indentation depth, and microtopographic variations, from a single exposure (Figure 2) [17]. Second, unlike multi-camera array systems, the MLA-based light-field camera employs a unified optical path and a single CMOS sensor, eliminating cross-device inconsistencies such as pixel shifts or geometric distortions. This substantially simplifies algorithmic pipelines by reducing the burden of system calibration, multi-view registration, and point-cloud fusion, thereby improving reconstruction robustness. Third, potentially integrating hyperspectral sensing will enable simultaneous acquisition of 3D surface geometry and internal biochemical information (e.g., moisture, protein, or pigment distribution). The joint analysis of spectral signatures and 3D morphology provides richer descriptors for predicting germination rate, seedling vigor, and other physiological traits, strengthening phenotype–genotype association studies [18,19].

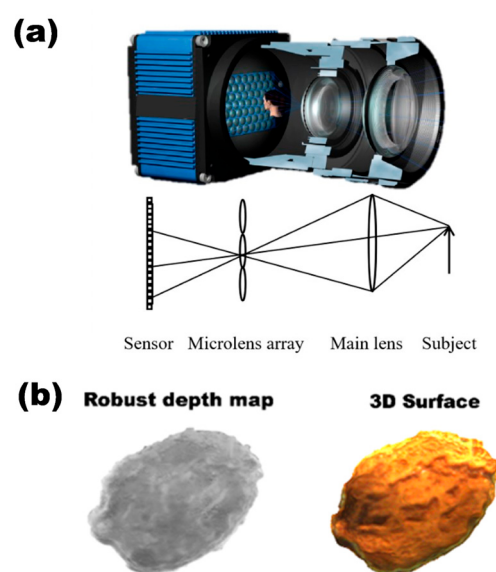


Figure 2. (a) Schematic of a microlens-array light-field imaging system. (b) Bitter melon 3D reconstruction surface [20].

Collectively, the metasurface-based light-field framework improves hardware compactness, has the potential to increase data dimensionality, and enhances interpretability, making it a compelling direction for next-generation seed phenotyping.

While metalens-array light-field imaging offers unique advantages, traditional MLA fabrication remains costly and technically challenging, especially when high layout precision or customized aberration control is required. Metasurface optics, consisting of subwavelength nano-resonators arranged on an ultrathin substrate, provide a transformative alternative. Metasurfaces feature nanoscale control over phase, amplitude, and polarization, and they are highly compatible with CMOS processes, making them a hallmark of next-generation optical components [21].

As illustrated in Figure 3, the achromatic metalens array developed by Fan et al. enables broadband 3D imaging across the visible spectrum (430–780 nm) via precisely engineered nanopillars. The 60×60 polarization-insensitive array demonstrates high efficiency (avg. 47%) and nearly diffraction-limited focusing while maintaining consistent focal planes. Experimental results show clear reconstruction of 3D scenes with accurate depth control under white-light illumination. This compact platform enables high-fidelity 3D scene reconstruction and represents a significant advancement for integral imaging, light-field cameras, and augmented reality applications.

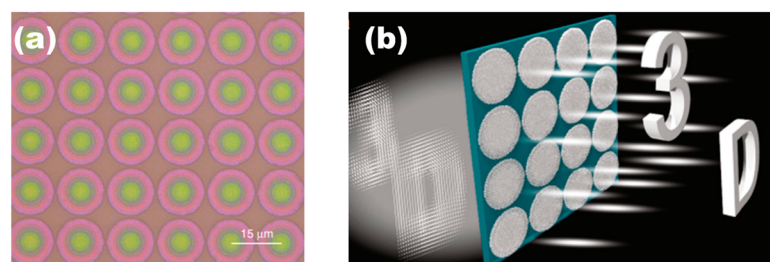


Figure 3. (a) SEM photo of achromatic metalens array and (b) metalens array used for 3D imaging [22].

Nippon Telegraph and Telephone Corporation (NTT) reported that a hyperspectral metasurface utilizing dispersion engineering could perform spectral encoding directly from compressed computation. NTT plans to mass-produce the technology for the robot sensing and smart agriculture markets. As illustrated in Figure 4, the metasurface array generates wavelength-dependent multiplexed images from a single exposure. A dedicated neural reconstruction network then jointly recovers a hyperspectral cube, effectively unifying previously separate imaging modules.

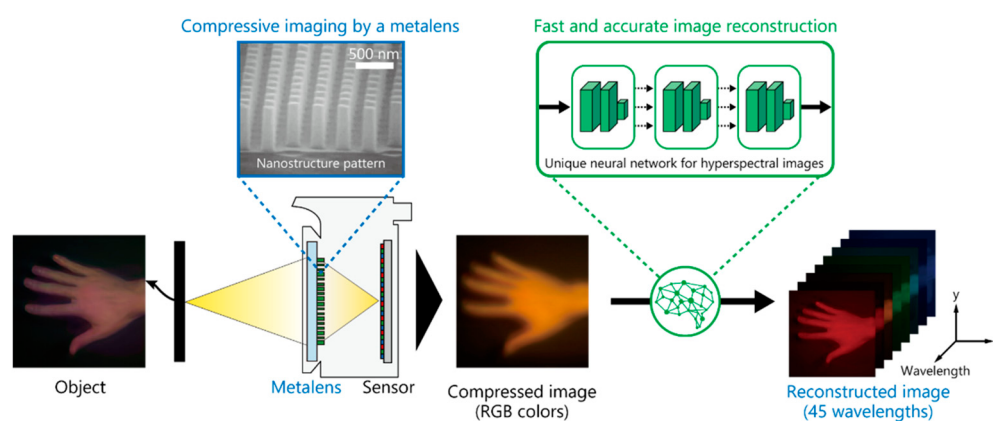


Figure 4. Schematic of the reconstruction process for hyperspectral metasurface imaging [23].

In summary, the metasurface-enabled light-field architecture delivers substantial gains in system compactness, optical efficiency, and multimodal data richness. By integrating 3D geometry and hyperspectral biochemical information into a single compact platform, the technology demonstrates significant potential for high-throughput, multimodal seed phenotyping and broader applications in precision agriculture, robotic vision, and biomedical imaging. In Section 2, we review recent advances in metasurface-lens-array light-field systems for hyperspectral imaging and 3D reconstruction, highlighting the underlying optical principles, design methodologies, and computational algorithms that enable single-shot, high-dimensional acquisition. We further outline an integrated system architecture, an application roadmap for seed phenomics, and the key technical challenges that shape the path toward a fully metasurface-enabled 3D–hyperspectral light-field imaging paradigm.

1.3. Scope and Contributions

To ensure a rigorous and comprehensive overview, we adopted a systematic search strategy targeting the intersection of three core domains: metasurfaces, computational imaging, and seed phenomics. We prioritized the literature bridging at least two of these fields, focusing primarily on the period from 2015 to 2025 to capture the rapid maturation of metasurface optics and deep learning, while also incorporating foundational milestone studies prior to this window to establish the necessary theoretical baselines for spatial–angular sampling principles.

Building on this selection, this review provides a unique contribution by explicitly mapping the convergence of these isolated fields, bridging critical gaps identified in current surveys. While current mainstream phenotyping platforms, such as conveyor-belt systems or gantry-based field scanners, are often limited by their bulky size, high calibration burden, and significant operational costs [24,25], we explore the transformative potential of compact metasurface; conversely, while existing photonics and computational surveys rigorously detail unit-cell physics [26] or light-field reviews [27,28], they often overlook specific biological integration challenges and spectral–spatial coupling. Hyper-spectral imaging surveys [29,30] predominantly discuss traditional scanning architectures or remote sensing applications, lacking coverage of emerging snapshot metasurfaces. By synthesizing these perspectives, we present the first integrated roadmap for a 5D (spatial–angular–spectral) imaging platform, offering a novel physics-informed coupling model to guide next-generation smart seed phenotyping.

2. Technical Foundations: From Light-Field Imaging to Metasurface Arrays

2.1. Principles of Light-Field Imaging and Metasurface Optics

The principle of light-field imaging lies in capturing both the spatial distribution and angular direction of light rays within a single exposure, thereby enabling the reconstruction of a three-dimensional scene. This approach is grounded in the four-dimensional light-field parameterization introduced by Levoy et al. [31], and its physical realization typically involves placing a microlens array in front of the image sensor in a conventional imaging system. As illustrated in Figure 5, each microlens partitions the exit pupil of the main lens into multiple discrete viewing directions, forming an array of micro-images on the sensor. By decoding the disparity among these micro-images, depth information can be inferred and the 3D structure can be reconstructed.

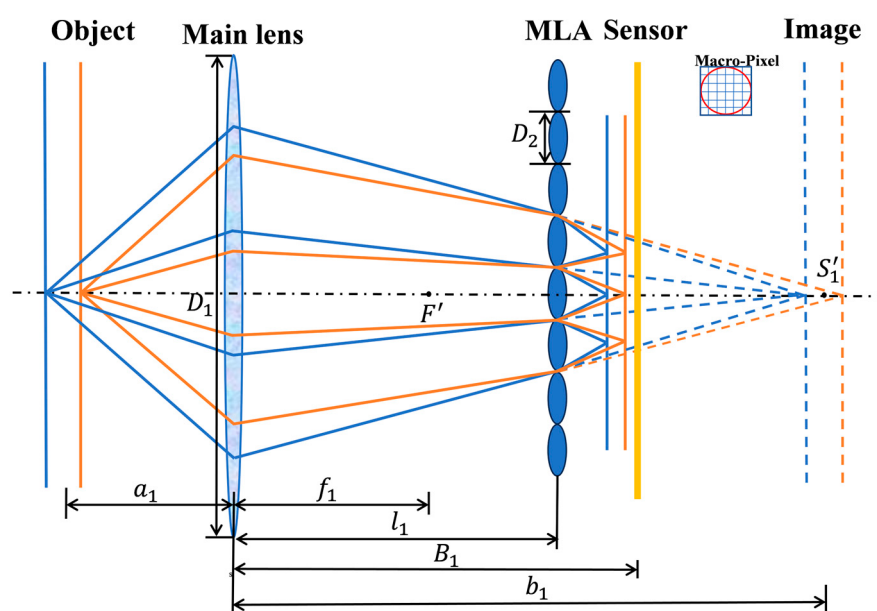


Figure 5. Optical path diagram of a focused light-field system (Galilean type) and schematic of macro-pixels [17].

As the mathematical foundation of light-field imaging, the light-field function $P(\theta, \varphi, \lambda, t, x, y, z)$ encapsulates the full seven-dimensional radiometric information carried by rays propagating through 3D space. In a focused light-field camera, the coordinated design of the microlens array and CMOS sensor realizes a structured four-dimensional

sampling of this function, yielding a discretized light-field matrix $L(x, y, u, v)$. The underlying mechanism relies on the wavefront-partitioning capability of the microlenses: each microlens maps rays from different incidence angles onto corresponding sub-aperture regions of the sensor, forming spatial–angular correlated sub-images.

Despite its versatility, traditional light-field imaging faces inherent technical bottlenecks imposed by its physical architecture. The most fundamental limitation arises from the trade-off in spatial–angular bandwidth: acquiring angular information inevitably sacrifices spatial resolution. Under typical imaging conditions ($\lambda = 488$ nm, $\text{NA} = 1.2$), the effective spatial resolution generally does not exceed ~ 3 μm , which is insufficient for high-fidelity microscopic phenotyping [32]. Moreover, insufficient angular sampling leads to pronounced axial aliasing, significantly degrading axial localization accuracy in conventional light-field systems [33].

To overcome the aforementioned limitations, metasurface-based flat optics have emerged as a revolutionary alternative to traditional refractive architectures. A metasurface consists of a two-dimensional array of subwavelength artificial “meta-atoms” arranged with precise spatial order. By engineering the geometric parameters of these nanostructures, arbitrary manipulation of the wavefront, including amplitude, polarization, and phase, can be achieved within an ultrathin, subwavelength thickness [34]. The phase profile of an ideal spherical lens can be expressed as

$$\phi(x, y) = \frac{2\pi}{\lambda} \left(\sqrt{f^2 + x^2 + y^2} - f \right) \quad (1)$$

where f is the focal length and λ is the operating wavelength. By arranging nanostructures with varying parameters at designated positions to provide full 2π phase coverage, an incident plane wave can be transformed into a spherical wave for focusing.

This design requires a pre-established library linking structural parameters to the optical phase response. Phase modulation in such metasurfaces is generally realized through one or a combination of three fundamental mechanisms: resonant phase, propagation phase, and geometric phase.

Resonant-phase modulation typically relies on electromagnetic resonance effects such as localized surface plasmon resonance (LSPR) in metallic structures, where a strong phase shift occurs when the incident wavelength matches the resonant frequency of the subwavelength elements. However, due to ohmic losses in metals and the narrowband nature of plasmonic resonances, such metasurfaces usually exhibit low transmission efficiency and limited bandwidth. Recently, all-dielectric metasurfaces have been widely adopted for resonant phase control, offering higher efficiency and stronger field confinement thanks to their low loss and high permittivity.

Propagation-phase modulation primarily utilizes the waveguide effect in dielectric nanostructures. For a nanopillar with a fixed height H , varying its transverse dimension (e.g., diameter) tunes the effective refractive index n_{eff} of the guided mode, thereby enabling controlled phase accumulation according to

$$\varphi = n_{\text{eff}} \frac{2\pi}{\lambda} H \quad (2)$$

Propagation-phase modulation exhibits polarization independence, while its phase accumulation correlates with the wavelength and effective refractive index. Intrinsically, it introduces material and structural dispersion, resulting in limited operating bandwidth. However, employing low-loss dielectric materials in the visible band enables high transmittance and high focusing efficiency.

The geometric phase (Pancharatnam–Berry phase) imparts a polarization-dependent phase profile ($\phi = 2\sigma\theta$, where $\sigma = \pm 1$ denotes the circular polarization handedness) by rotating anisotropic meta-atoms [35]. It operates via spin–orbit interaction, offering inherent dispersion-free operation. This mechanism is intrinsically polarization-selective: only the spin component matching the meta-atom’s handedness is efficiently diffracted, while the orthogonal spin is filtered. However, the absence of dispersion in the geometric phase itself limits its direct use as a dispersive control element in applications such as hyperspectral imaging, where wavelength-dependent phase manipulation is often required. Moreover, when implemented using a half-wave retarder, its theoretical diffraction efficiency is limited to 50% owing to the conjugate conversion between the two circular polarization states.

Currently, to achieve broadband dispersion and efficient spectral imaging and reconstruction, next-generation smart seed phenotyping platforms are increasingly adopting all-dielectric metasurfaces that utilize composite phase-modulation mechanisms [36]. These platforms employ inverse design methods to co-optimize the geometric parameters (such as shape and size) and spatial arrangement of the meta-atoms. This hybrid phase-modulation strategy introduces additional degrees of design freedom, enabling effective compensation or customization of the device’s dispersion characteristics. Consequently, it achieves stable phase responses across a broad spectral range and ultimately enhances the reconstruction quality and accuracy of hyperspectral data. Figure 6a illustrates the designed phase map of a spherical metalens together with its simulated focusing performance [37]. Figure 6b shows a schematic of a metasurface unit cell, in which phase modulation is achieved by tuning structural parameters such as height and width. Figure 6c plots the transmission amplitude and phase response as functions of unit-cell width, while Figure 6d displays the design layout of a 10×10 metalens array.

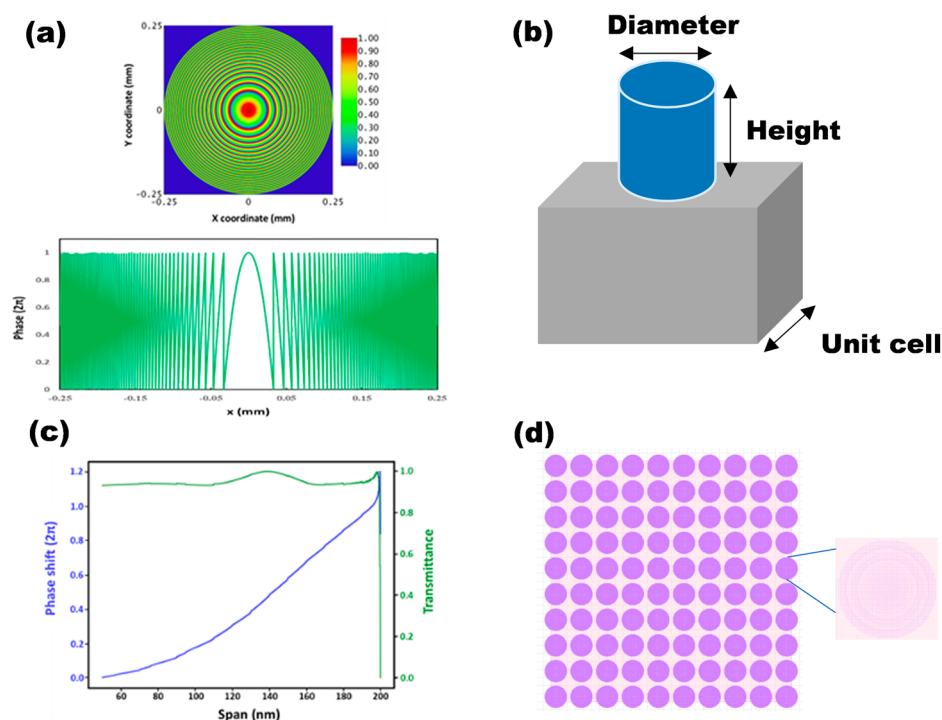


Figure 6. Metalens array design flow. (a) The ideal target phase profiles. (b) The schematic diagram of a subwavelength structure within a unit cell. (c) Data plots comparing phase shift, transmittance, and span. (d) A 10×10 metalens array and single metalens design layout [37].

Metalenses fabricated via nano-fabrication technology offer several advantages that are unattainable with conventional refractive optics. Their fabrication is fully compatible with

mature semiconductor CMOS processes, and the resulting devices are ultrathin, lightweight, highly integrable, and amenable to low-cost, mass production. Replacing traditional microlens arrays with metasurface lens arrays in light-field cameras has therefore been widely recognized as an effective solution to overcoming fabrication complexity, high cost, and inherent performance constraints of conventional microlens systems. This advancement lays the technological foundation for next-generation compact, high-performance light-field imaging platforms.

Metasurface technology is widely regarded as one of the most promising approaches for advancing light-field imaging. Lin's group at Nanjing University applied achromatic metalens techniques to light-field camera design [38], successfully realizing optical digital zoom functionality. As shown in Figure 7, the achromatic metalens-array (AMLA) light-field camera modulates the incident light field through the metasurface array and forms a light-field image on the sensor. Reconstruction algorithms then refocus the captured light field computationally, generating digitally zoomed images at different focal depths. This demonstration highlights that metasurface arrays provide an effective solution to the long-standing challenges of microlens arrays, namely fabrication complexity and high cost, while simultaneously enabling system-level integration of light-field cameras. Fan et al. demonstrated a trilobite-inspired metalens array that achieves an extreme DOF ranging from centimeters to kilometers, a capability crucial for imaging multi-layer seed piles without refocusing [39]. Furthermore, while MLAs are polarization-blind, Capasso exploited the birefringence of meta-atoms to realize a compact, single-shot polarization imaging system [40]. Similarly, regarding 3D displacement and strain measurement, Zhao et al. [41] recently demonstrated a compact binocular metalens camera. By integrating two GaN-based metalenses on a single substrate, this system achieved high-precision 3D Digital Image Correlation (DIC) with a baseline of only 4 mm, offering a lightweight solution for monitoring surface deformation. This allows for the simultaneous capture of 3D geometry and surface texture anisotropy, providing a new dimension of phenotypic data that traditional refractive optics cannot access. Song et al. demonstrated a varifocal meta-device for augmented reality that dynamically tunes the focal plane without mechanical movement [42], while Zhou et al. introduced a tunable point-cloud projection meta-device capable of high-resolution 3D structured light imaging. These innovations signal a shift toward active metasurface systems [43].

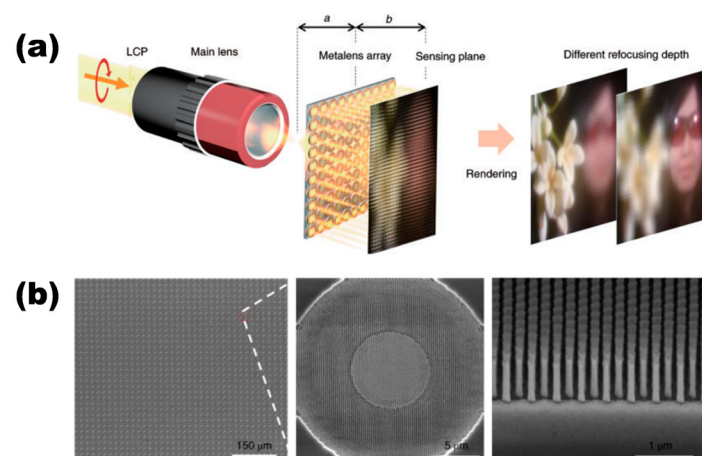


Figure 7. (a) Schematic of achromatic metalens-array (AMLA) light-field depth imaging and (b) SEM images of metasurface array, single metalens, and zoomed metalens [38].

2.2. Dispersion Engineering of Metasurfaces and Hyperspectral Imaging Mechanisms

Unlike conventional refractive optics, which rely on accumulated phase delay through propagation, metalenses achieve wavefront shaping via localized resonances and geometric-phase mechanisms. Each subwavelength meta-atom must satisfy a local phase-matching condition, expressed in Equation (2):

$$\varphi(x, y) = -\frac{2\pi}{\lambda_d} \left(\sqrt{(x - x_f)^2 + (y - y_f)^2 + z_f^2} - f \right) \quad (3)$$

where $f = \sqrt{x_f^2 + y_f^2 + z_f^2}$ denotes the focal length, and φ and λ_d represent the target phase and design wavelength, respectively.

Because the phase distribution of a metalens varies with incident wavelength, dispersion inevitably induces focal shifts. In coaxial optical configurations, the rotational symmetry of the phase map causes focal spots of different wavelengths to shift longitudinally along the optical axis. For hyperspectral imaging, the spectral resolution $\delta\lambda$ is a key performance metric. The focal length f , focusing angle α , and numerical aperture (NA) are the primary factors governing spectral resolution. From the Rayleigh criterion, $0.61\lambda/NA$, the minimum resolvable wavelength interval in the focal plane can be determined. The spectral resolution is given by Equation (3):

$$\delta\lambda = \frac{\Delta\lambda}{f \left\{ \sin^{-1} \left[\left(1 + \frac{\Delta\lambda}{\lambda_d} \right) \sin(\alpha) \right] - \alpha \right\}} \times \frac{0.61\lambda}{NA} \quad (4)$$

This wavelength-dependent focusing behavior enables metalens arrays to integrate functionalities such as multi-focus operation and achromatic correction within sub-millimeter thicknesses, significantly enhancing system compactness [44]. Nevertheless, dispersion and angular sensitivity of subwavelength resonators impose constraints on efficiency: typical transmission efficiencies (<60%) remain considerably lower than those of refractive lens arrays (~90%), constituting a central bottleneck for metasurface-based imaging.

Based on their imaging mechanism, hyperspectral metasurface cameras can be classified into transverse-dispersion and longitudinal-dispersion designs. Longitudinal-dispersion systems require only a single CMOS sensor and offer superior compactness. However, because images from all wavelength channels are spatially overlapped, they demand highly sophisticated reconstruction algorithms. This challenge makes their implementation more difficult than transverse-dispersion spectrometers, and consequently, fewer studies have been reported in this category.

Axial chromatic dispersion arises because light at different wavelengths experiences different refractive indices when passing through an optical element, leading to focal shifts along the optical axis. As illustrated in Figure 8a, this behavior is inherent to conventional refractive lenses. Metalenses exhibit similar wavelength-dependent focal shifts because their phase distribution, governed by Equation (5), varies with incident frequency (Figure 8b). This chromatic dispersion degrades image sharpness, also motivating achromatic metalens designs such as those shown in Figure 8c [45].

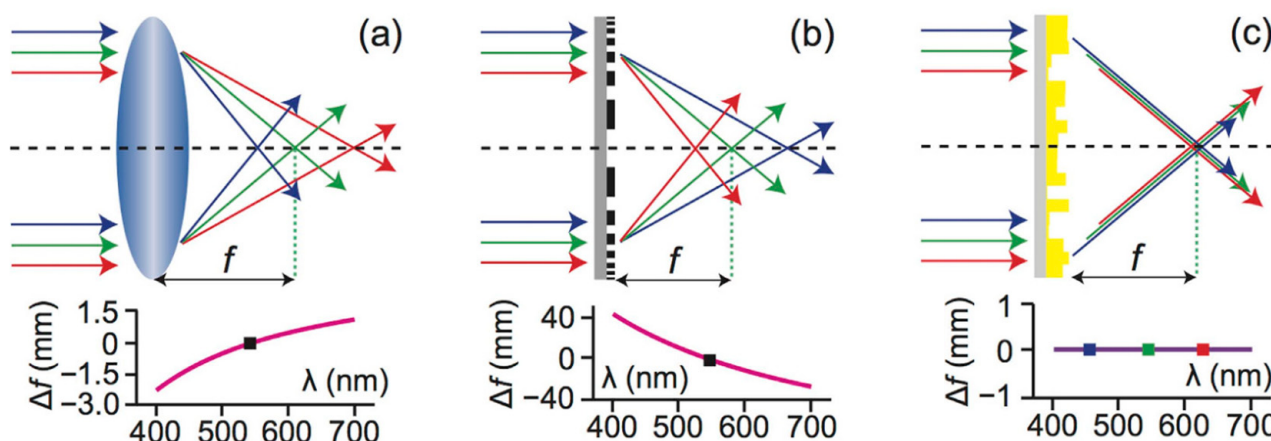


Figure 8. Axial chromatic dispersion in (a) conventional refractive lenses, (b) dispersive metalenses, (c) achromatic metalenses. The upper panels illustrate focal-length variation with wavelength [45].

Dispersion in metasurfaces can be actively engineered to serve as a spectral encoding resource. Through inverse design, the group delay dispersion within the phase response is precisely tailored to establish a strong, controllable linear dependence between the focal length and the optical frequency. This engineered chromatic aberration, dominated by group delay dispersion, causes different frequencies to form spatially separated point-spread functions (PSFs) at the image plane—each uniquely associated with a specific wavelength. As a result, the spectral information of the incident light is directly encoded into the spatial domain. To obtain the group-delay response, the phase distribution in Equation (5) is expanded in a Taylor series around the design angular frequency ω_d , typically chosen as the center frequency of the operational bandwidth:

$$\varphi(r, \omega) = \varphi(r, \omega_d) + \left. \frac{\partial \varphi(r, \omega)}{\partial \omega} \right|_{\omega = \omega_d} (\omega - \omega_d) + \frac{1}{2} \left. \frac{\partial^2 \varphi(r, \omega)}{\partial \omega^2} \right|_{\omega = \omega_d} (\omega - \omega_d)^2 + \dots \quad (5)$$

Here, $\varphi(r, \omega_d)$ describes the ideal phase profile at the center frequency; the first derivative term corresponds to the group delay, while the second derivative term captures group-delay dispersion. Under broadband illumination, the incident field can be treated as a superposition of independent frequency components. In achromatic metalenses, the target phase $\varphi(r, \omega_d)$ steers all frequency components toward the same focal point, whereas the engineered group delay compensates the interband phase mismatches through controlled temporal offsets. This dual-control mechanism enables metalenses to maintain stable, focusing performance across a broad spectral range. Conversely, the same dispersion mechanism can be exploited for hyperspectral imaging using metasurfaces [46]. Metalenses with engineered dispersion are emerging as a powerful paradigm for computational imaging. By design, they spatially separate different wavelengths, allowing hyperspectral images to be retrieved mathematically using prior knowledge. Beyond imaging, dispersion engineering is also revolutionizing precision metrology. As illustrated in Figure 9, He et al. recently proposed a dispersive metalens thermometry system that encodes spectral temperature information into compressive measurements, achieving high-accuracy high-temperature sensing via deep learning decoding [47].

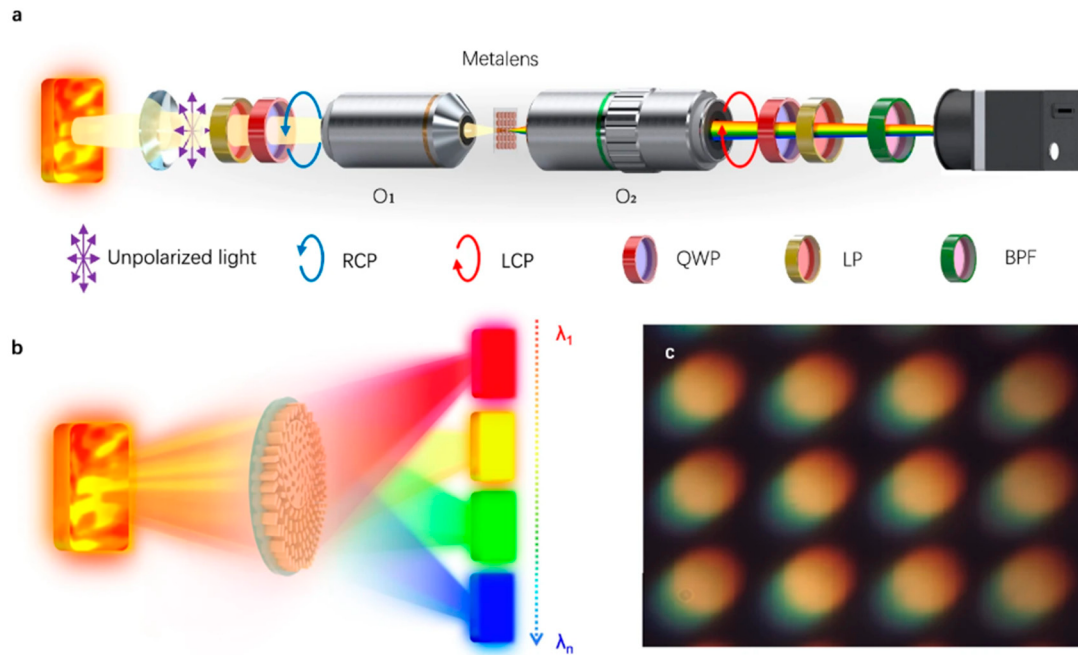


Figure 9. Experimental setup of dispersive metalens thermometry. (a) Three-dimensional diagram of the DMT experimental setup. (b) Schematic illustration of the dispersion mode of the metalens. (c) Raw dispersion image of blackbody furnace at 1873 K [47].

Tan et al. proposed a metasurface-based imaging framework that integrates a multi-depth, multiwavelength-focusing metalens with an RGB reconstruction neural network to recover depth and color information from intentionally defocused images [48]. Simulations of a 1 mm diameter metalens demonstrate that the system can capture 3D depth and texture information over a working range of 0.12–0.6 m. The point-spread functions (PSFs) corresponding to different focal depths and wavelengths are first numerically modeled. The metasurface’s modulation of the incident optical field is expressed as

$$T_{\lambda}(x, y) = A(x, y) \sqrt{\text{PCE}_{\lambda}(x, y)} \exp(j\Phi_{\lambda}(x, y)) \quad (6)$$

where $A(x, y)$ is the circular aperture function, $\sqrt{\text{PCE}_{\lambda}(x, y)}$ represents amplitude modulation induced by the metasurface, and $\Phi_{\lambda}(x, y)$ denotes its phase modulation. Under the Fresnel approximation, the PSF can be written as

$$P_{\lambda, z}(x', y') \propto \left| \mathcal{F} \left\{ A(x, y) \sqrt{\text{PCE}_{\lambda}(x, y)} \exp \left\{ j \left[\frac{\pi}{\lambda} \left(\frac{1}{z} + \frac{1}{z'} \right) (x^2 + y^2) \right] \right\} \right\} \right|^2 \quad (7)$$

where z' is the distance from the metasurface to the detector and (x', y') denote detector-plane coordinates. Based on these single-wavelength PSFs, Tan’s group proposes constructing a deconvolution-based neural network that leverages PSFs across the wavelength dimension to synthesize continuous spectral cross-sections. By combining physical priors with data-driven training on large image datasets, the network learns wavelength-dependent feature representations and subsequently reconstructs high-fidelity hyperspectral images for each channel, as outlined in Figure 10 [49].

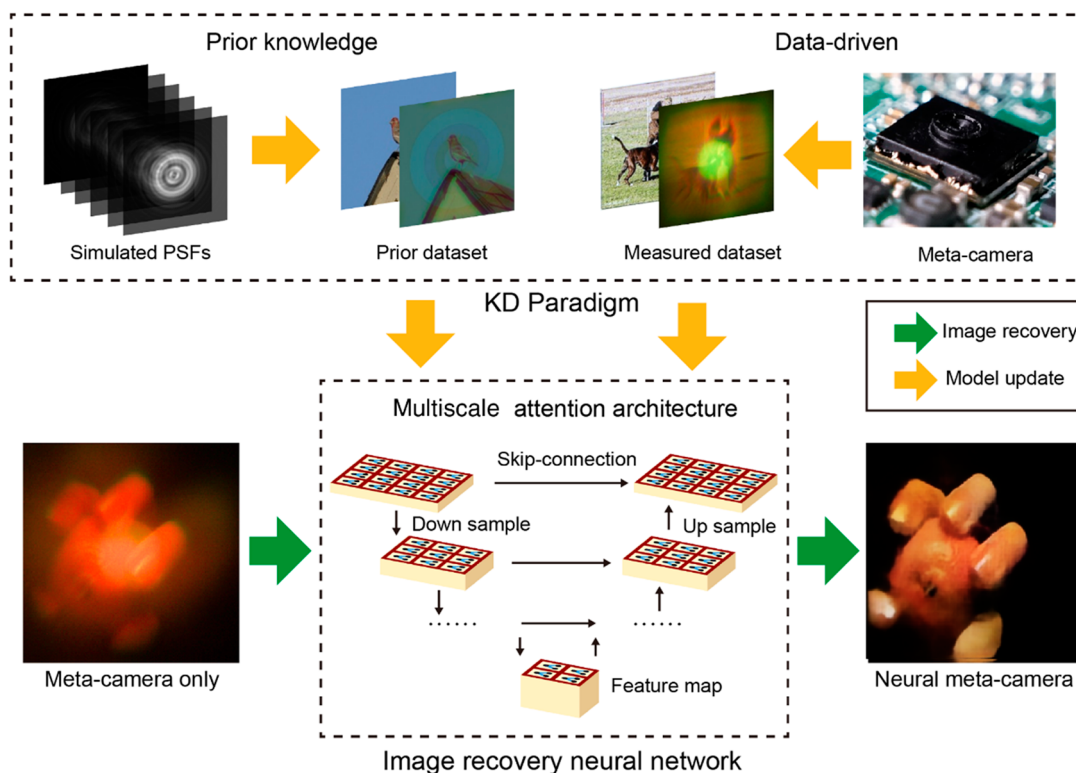


Figure 10. Reconstruction pipeline for hyperspectral imaging using a dispersive metalens [49].

Hua et al. at Nanjing University demonstrated an ultra-compact spectral light-field imaging architecture, SLIM, based on a laterally dispersive metasurface array and a monochrome sensor [50]. A single-shot SLIM acquisition achieves a spectral resolution of 4 nm with near-diffraction-limited spatial resolution. Figure 11a,b show the schematic of the lateral-dispersion imaging system and corresponding SEM images of the metasurface, while Figure 11c presents a raw laterally dispersed image. A comparison between the SLIM metasurface-based system and a traditional light-field camera is provided in Figure 12.

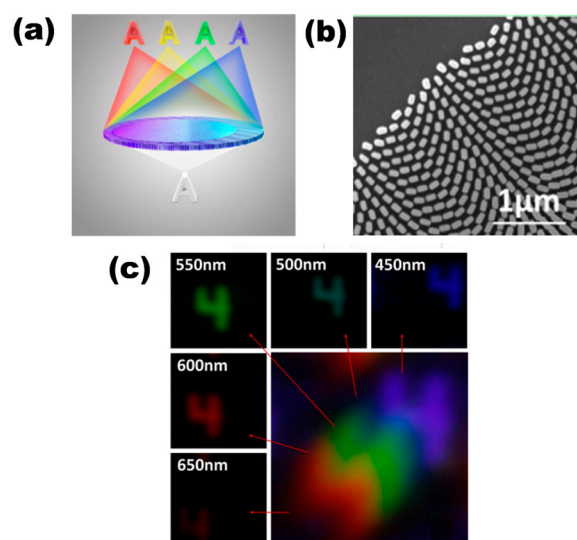


Figure 11. (a) A schematic of the lateral-dispersion imaging system; (b) corresponding SEM images of the metasurface. (c) A raw laterally dispersed image [50].

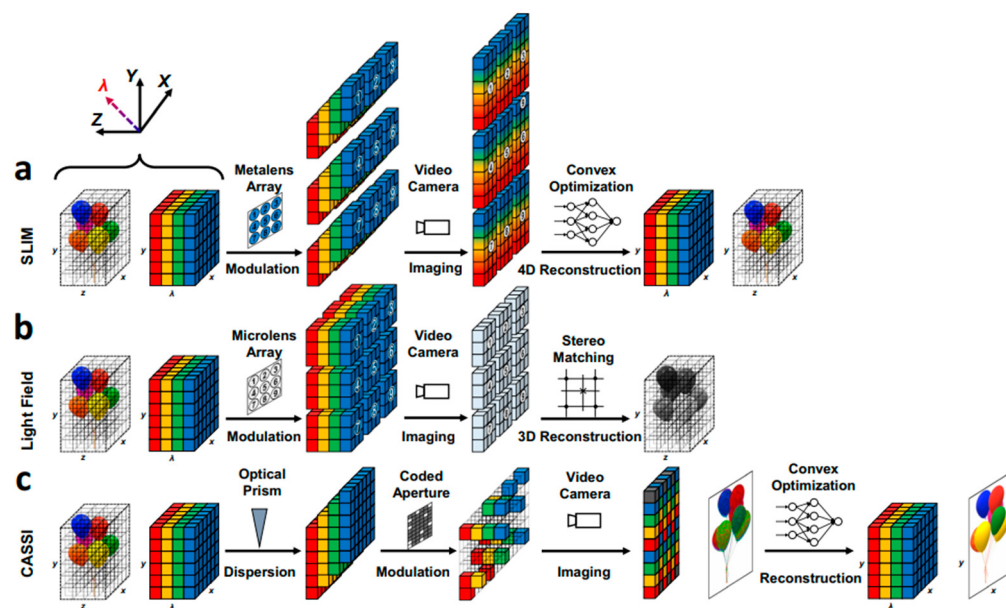


Figure 12. SLIM spectral light-field imaging system and benchmarking. Schematic workflow of (a) SLIM system, (b) light-field system, (c) Coded Aperture Snapshot Spectral Imaging (CASSI).

The SLIM architecture exploits lateral dispersion to map different wavelengths onto laterally shifted positions, producing a “motion-blur-like” encoded measurement in which spectral information is implicitly embedded. Using the dispersion model of the metasurface, the system generates a laterally blurred forward image, from which a spectral reconstruction algorithm retrieves a 21-channel hyperspectral data cube. Representative reconstruction results are shown in Figure 13.

The field has recently witnessed a paradigm shift from optimizing individual components to end-to-end system design, an emerging frontier recently synthesized and theoretically framed as “computational metaoptics” by Carnes [51]. Foundational to this new paradigm are breakthroughs in hardware efficiency: for instance, Chen demonstrated dispersion-engineered metasurfaces with over 90% relative diffraction efficiency across the visible spectrum, effectively mitigating the low-signal-to-noise-ratio issues that previously hindered computational reconstruction [52]. Building on such high-performance hardware, Heidrich’s group advanced the system architecture by introducing collaborative metalens arrays where millions of nanopillars are inversely designed to produce complementary point-spread functions [53]. To further ensure robustness in wide-field imaging, critical for curved seed surfaces, Cui. proposed topology-optimized metasurfaces that maintain spectral consistency even under oblique incidence [54]. Additionally, Bao et al. validated the capability of metasurfaces for single-shot, multidimensional (intensity, phase, and polarization) information capture [55].

Overall, metasurfaces with engineered dispersion profiles provide a promising route toward compact multispectral and hyperspectral imaging. However, lateral-dispersion architectures inherently require spatially separated detector sub-arrays, which increase system footprint, design complexity, and cost. In contrast, the axial-dispersion configuration enables a substantial reduction in system volume and hardware complexity. This advantage, however, comes at the expense of significantly more challenging spectral reconstruction, necessitating innovation in dispersion-engineered metalens design, system-level architecture, and learning-based reconstruction algorithms.

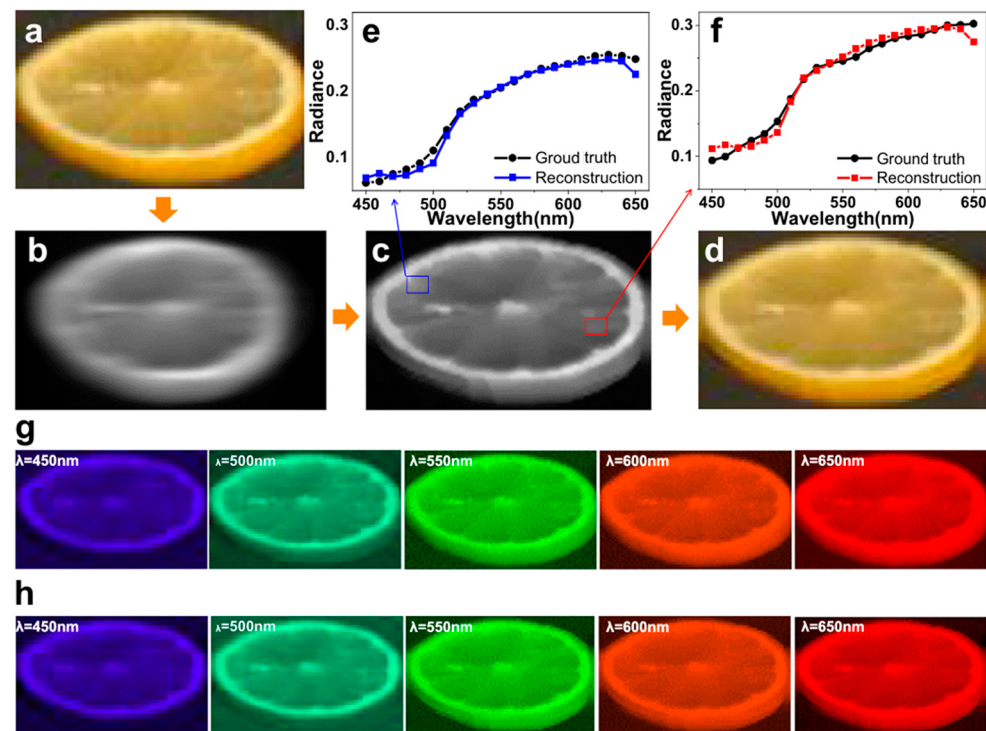


Figure 13. Imaging demonstration of SLIM spectral reconstruction. (a) Ground-truth RGB image; (b) simulated dispersive image using the SLIM forward model; (c) reconstructed grayscale image; (d) synthesized RGB image from reconstructed spectra; (e,f) spectral profiles at two spatial locations comparing reconstruction (colored) with ground truth (black); (g,h) raw and reconstructed monochromatic images for selected wavelengths.

3. Design and Optimization of Light-Field Modulation by Metasurface Lens Arrays

3.1. Architecture for Metasurface-Based Light-Field Manipulation

In Section 2, we discussed the principles of light-field imaging and the wavelength-dependent wavefront control by metalenses, providing examples of 3D depth sensing and snapshot multispectral imaging achieved through both lateral and axial dispersion engineering. In this section, we discuss the principle by which the metalens manipulates multidimensional light-field modulation via metalens arrays, and the design and optimization philosophy required to implement large, efficient metalens and metasurface devices in practice.

In order to implement metasurface-based light-field devices with hyperspectral imaging capability, both optical hardware and software need to be investigated innovatively [56]. The schematic in Figure 14 illustrates the workflow of implementing a multidimensional light-field camera based on a metasurface lens array as the key dispersive element. This workflow encompasses several critical technical challenges:

1. **Multidimensional light-field design:** This requires leveraging dispersion theory, aberration analysis, and related optical design principles to guide the design and optimization of metalenses [57]. A comprehensive library of nanostructured pillar units is established, along with a complete dispersion phase model. Inverse-design methods are then employed to jointly optimize the metasurface lens array and hyperspectral performance [58].
2. **Fabrication of large-aperture metasurface arrays:** The design specifications necessitate advanced fabrication processes capable of producing large-aperture metalens arrays [59]. While DUV steppers offer high resolution for subwavelength structures,

their exposure field is typically limited. Fabricating a large-aperture metalens array exceeding this field requires reticle stitching. Technical challenges here include controlling stitching errors at the nanometer scale; slight misalignments at exposure boundaries introduce phase discontinuities, leading to scattering and ghost images that degrade the reconstruction quality of the light field. Furthermore, the high cost of DUV photomasks and the low throughput associated with multi-shot exposures pose a barrier to cost-sensitive agricultural applications. Optimization of fabrication workflows and post-fabrication characterization of large-aperture samples are critical to ensure performance consistency. To address the cost constraints, NIL is viewed as a promising alternative for mass production. However, for high-aspect-ratio metasurface pillars, demolding defects (where nanopillars fracture or adhere to the mold) significantly reduce yield. A critical barrier in large-area NIL is maintaining the uniformity of the residual layer thickness across the entire wafer. Variations in RLT alter the effective height of the meta-atoms, causing phase errors that result in chromatic aberration, which is particularly detrimental for hyperspectral reconstruction accuracy. Despite these hurdles, recent progress has been made. Rho's group recently demonstrated a "roll-to-plate" printing protocol capable of fabricating centimeter-scale RGB achromatic metalenses, significantly lowering the unit cost for mass deployment [60].

3. Multidimensional light-field reconstruction: Accurate reconstruction requires knowledge of the point-spread function (PSF) of the metasurface array. Using broadband spectral characteristics, deconvolution and end-to-end neural network models are developed to reconstruct hyperspectral images [61]. This reconstruction framework is possibly applied to seed phenotyping experiments, ultimately enabling the creation of a 3D hyperspectral imaging database for seeds.

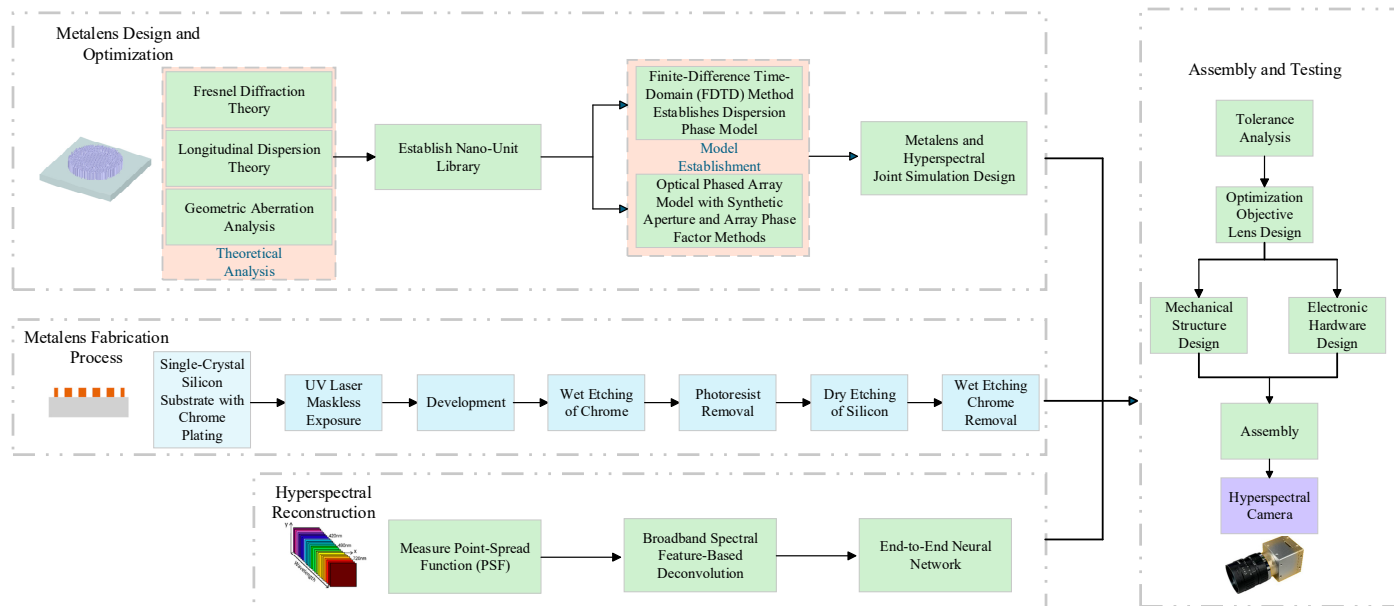


Figure 14. Technical implementation workflow for hyperspectral light-field imaging using a metasurface lens array.

3.2. High-Efficiency Inverse Design of Metasurfaces

Current large-area metasurfaces face several key limitations that constrain optical efficiency. First, chromatic and monochromatic aberrations impose inherent trade-offs between focusing efficiency and operational bandwidth, representing the principal factor restricting metasurface performance. Second, the design of metasurfaces is intrinsically a global optimization problem with a high-dimensional parameter space. The lack of

efficient optimization algorithms often prevents achieving optimal device performance. Third, fabrication of metasurface unit cells demands feature accuracy on the order of $\lambda/5$ to $\lambda/10$. Meeting such stringent precision requirements necessitates high-end lithography equipment and novel process development, which both increases the technological barrier for widespread adoption and elevates manufacturing costs.

Most conventional metasurface designs follow a forward-design approach, in which a unit-cell library is first constructed through parameter sweeps, and then individual unit geometries are arranged according to a pre-determined phase profile. Forward design performs adequately for simple functionalities, such as small-aperture, single-wavelength focusing lenses or meta-gratings. However, as design complexity and constraints increase, forward design struggles to find globally optimal solutions. Realizing the next generation of metasurface devices therefore requires a paradigm shift in design philosophy.

Compared with forward design, inverse design is function-driven approach: target optical performance guides the optimization of individual unit geometries using advanced search algorithms. Originally developed for large-scale engineering optimization problems, such as bridge or aircraft wing shape design, inverse design has recently been applied to micro- and nanophotonic device design. Current approaches include gradient-based topology optimization and continuous contour shape optimization.

Capasso's group proposes a hybrid inverse-design method that combines genetic-algorithm-driven intelligent search with adjoint-field optimization [52]. Specifically, fast solver is widely used to shorten the simulation time, and backward-propagated fields are transmitted through the metasurface device. The optimal structural parameters are selected from a pre-built high-diffraction-efficiency unit library (efficiency $\sim 90\%$). Adjoint optimization accelerates convergence of large-aperture metasurface parameters, achieving globally optimal designs at minimal computational cost. Figure 15 illustrates the framework of this inverse-design process [62].

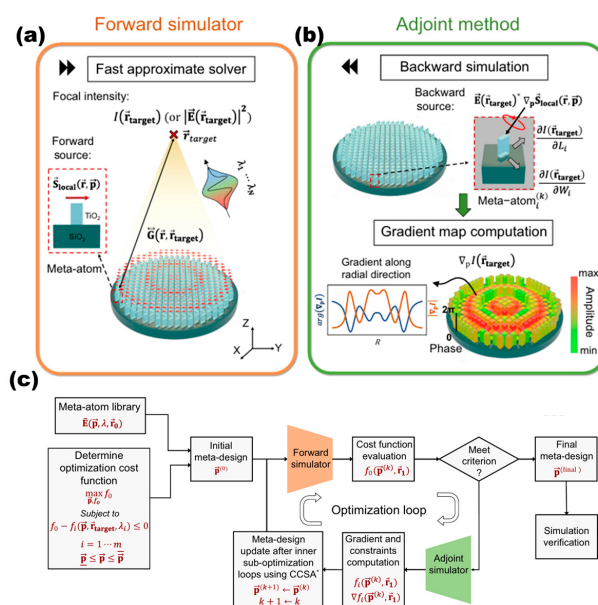


Figure 15. Efficient inverse-design framework for metasurfaces based on adjoint-field optimization. (a) Forward simulation of the metalens. (b) Rapid calculation of parameter sensitivities via adjoint fields. (c) Optimization workflow combining adjoint-field information with gradient-descent algorithms [52].

As the dimensionality of metasurface designs increases, simultaneously handling nanometer-scale unit cells (~ 10 nm) and macroscale devices ($100 \mu\text{m} \sim 10$ cm) becomes challenging. For instance, a $50 \mu\text{m}$ -diameter device with a 5 nm mesh resolution requires

~100 h of computation and ~100 GB of RAM. By integrating rigorous coupled-wave analysis (RCWA) with adjoint-based array optimization, it becomes feasible to design large-scale, high-performance metasurfaces in few hours with lower memory requirements. This inverse-design approach is computationally efficient and particularly advantageous for metasurface devices spanning nanoscale to macro-dimensions (>1000 wavelengths). Using this methodology, metasurfaces have been implemented with apertures ranging from micron-meters to centimeters, representing a four-orders-of-magnitude increase in area relative to existing technologies. These results demonstrate the feasibility of large-aperture, dispersive metasurface lens arrays for practical implementation [63].

4. Computational Reconstruction and Information Decoupling Algorithms

The reconstruction of high-dimensional data from compressive measurements is a foundational challenge in computational imaging. In this section, we review the established algorithms currently employed in the field, serving as the algorithmic building blocks for the integrated 5D imaging platform we propose in Section 5. Specifically, Section 4.1 establishes the theoretical projection models and resolution limits of metalens arrays. Subsequently, Section 4.2 surveys existing geometric and deep-learning methods for depth estimation, while Section 4.3 reviews neural network architectures for spectral decoupling.

4.1. Three-Dimensional Light-Field Analysis and Reconstruction from Metalens Arrays

A metalens-array light-field camera captures light signals from multiple directions in space by placing a metalens array in front of a CMOS sensor. During calibration and correction, combining this with the light-field camera projection model enables accurate retrieval of the light-field information.

After passing through the main lens, light is focused by the metalens array onto the CMOS sensor. Images from different viewpoints record rich spatial information, providing the data foundation for depth estimation and three-dimensional reconstruction. The light-field function $L(u, v, s, t)$ describes the intensity of rays passing through spatial coordinates (u, v) and arriving at (s, t) :

$$I_{u_i, v_j}(s, t) = L(u_i, v_j, s, t) \quad (8)$$

where (u_i, v_j) are fixed coordinates on the UV plane, corresponding to the viewpoint of the $i \times j$ -th sub-aperture image, and (s, t) are coordinates on the ST plane, corresponding to pixels within each sub-aperture image.

4.1.1. Spatial and Angular Resolution

For a focused light-field camera, the decoded spatial resolution of the raw light-field image is given by the sensor resolution multiplied by b/a , while the angular resolution is a/b . Here, b must satisfy

$$\frac{b}{a} = \begin{cases} 1 - \frac{b}{f_m}, & (0.5f_m < b < f_m) \\ \frac{b}{f_m} - 1, & (f_m < b < 1.5f_m) \end{cases} \quad (9)$$

where a is the distance from the main-lens image plane to the MLA, b is the distance from the MLA to the CMOS plane, and f_m is focus length of MLA. The spatial resolution of the light-field image increases as a decreases. Importantly, for focused light-field cameras, the trade-off between spatial and angular resolution is independent of the number of microlenses, allowing relatively large microlens units to mitigate edge effects [64,65].

4.1.2. Disparity Map Accuracy

The imaging quality and disparity-map precision of a light-field camera are jointly determined by

$$\Delta Z = \frac{\overline{Z^2} \cdot \Delta p}{d \cdot p_{\text{MLA}} \cdot f_{\text{MLA}}} \quad (10)$$

where ΔZ is the disparity-map precision, Z is the scene depth, Δp is the pixel size, d is the distance between the main lens and the microlens array, p_{MLA} is the microlens pitch, and f_{MLA} is the microlens focal length.

This provides a quick estimation for analyzing the spatial–angular resolution and depth precision of metalens-array light-field cameras, supporting high-fidelity 3D reconstruction for hyperspectral imaging and seed phenotyping applications [66].

4.1.3. Three-Dimensional Light-Field Reconstruction Algorithm

Light-field imaging captures multidimensional ray information of a scene, providing rich geometric and viewpoint cues for three-dimensional reconstruction. Traditional multi-view reconstruction methods rely on feature matching and dense correspondences. However, they are prone to ambiguity in regions with occlusion or weak texture, limiting reconstruction accuracy.

As illustrated by the workflow in Figure 16, Epipolar Plane Image (EPI)-based light-field depth estimation has become one of the mainstream approaches for achieving high-precision 3D reconstruction. The core idea is to extract 2D EPI slices from the light-field data volume, where pixels form linear structures whose slopes directly encode scene depth. By analyzing photoconsistency, texture orientation, or structure tensors within these EPIs, reliable disparity maps can be obtained, which are then converted to depth maps using pre-calibrated camera parameters and fused with the central-view RGB image to generate dense, full-color 3D point clouds.

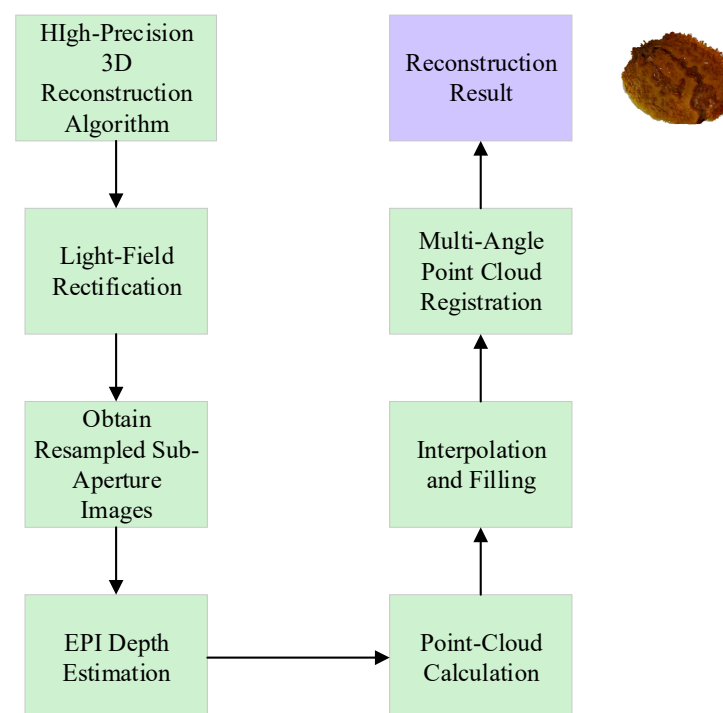


Figure 16. Workflow of the three-dimensional light-field reconstruction algorithm.

To enhance robustness on challenging surfaces such as seeds, with complex textures, low-texture regions, and arbitrary orientations, recent methods have commonly incor-

porated adaptive cost aggregation, occlusion-aware weighting, or multiscale analysis strategies. These improvements enable accurate depth estimation even in difficult areas with high curvature or specular reflections. EPI-based pipelines offer high computational efficiency, are easily integrated with deep-learning modules, and have been widely applied to agricultural light-field imaging tasks, including seed phenotyping. They provide a practical and effective 3D reconstruction solution for metasurface-enabled light-field systems.

4.2. Light-Field Depth Estimation Methods

Light-field depth estimation methods fall into two main categories: geometric approaches and deep learning-based techniques. Both exploit the fundamental property that scene points project onto lines in EPIs, where line slopes directly encode depth information.

4.2.1. Geometric Methods

Geometric methods build on the observation that depth estimation reduces to computing line orientations in EPI representations (Figure 17), utilizing photoconsistency constraints and the inherent structure of light-field data. Wanner and Goldluecke [67] pioneered a variational framework using structure tensor analysis on EPIs for local disparity estimation. The method achieves sub-pixel accuracy through continuous disparity space representation, circumventing the discretization problems common in traditional multi-view stereo. Local estimates are consolidated into globally consistent depth maps via convex optimization with ordering constraints that enforce physically plausible occlusion relationships. Their denoising scheme maintains competitive accuracy while achieving real-time performance, demonstrating that continuous disparity representation enables both precision and efficiency.

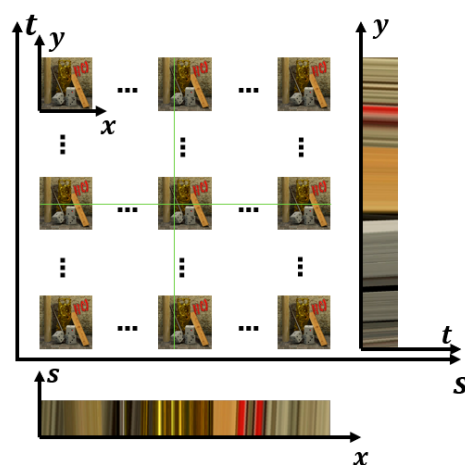


Figure 17. Principle of EPI-based depth estimation.

Building on this foundation, Zhang et al. introduced the spinning parallelogram operator (SPO) to better handle occlusion and noise [68]. The SPO exploits the principle that pixel intensity distributions on opposite sides of the correct depth line should maximally differ. The parallelogram orientation indicates depth, determined by maximizing the χ^2 distance between the two distributions. This formulation proves particularly robust because it maintains correct depth information even when matching ambiguities would confound conventional methods. Confidence metrics integrate redundant information across horizontal, vertical, and diagonal EPIs, with weighting functions accounting for distance from the center line to handle cross-talk artifacts and low angular resolution characteristic of plenoptic cameras. The method demonstrates superior performance near

occlusion boundaries while remaining insensitive to angular resolution constraints, making it especially suitable for consumer-grade devices.

For plenoptic cameras specifically, Sabater (2015) developed a pipeline emphasizing proper demultiplexing without preliminary demosaicing to prevent cross-talk artifacts [69]. Their block-matching algorithm exploits the natural rectification and uniform baseline of plenoptic camera views, simplifying the correspondence problem. A key contribution is the constraint that horizontal and vertical disparities are equal at each point, a consequence of the circular microlens array symmetry. This constraint improves matching accuracy for hexagonal microlens arrangements in Lytro-type cameras, where standard Bayer pattern assumptions break down.

Wang advanced occlusion handling by observing that approximately half the angular patch remains photoconsistent at correct depth, even for occluded pixels [70]. Their framework partitions angular patches into all regions based on geometric relationships between occlusion edge orientation in the spatial domain and the line separating consistent and inconsistent regions in the angular domain. Computing separate variances and analyzing their ratio enables occlusion likelihood estimation, focusing depth computation on reliable unoccluded portions. This approach yields substantial improvements at depth discontinuities, achieving sharper object boundaries and more accurate depth transitions where traditional photoconsistency assumptions fail.

4.2.2. Deep Learning Methods

Deep learning approaches learn the depth estimation in the form of “end-to-end” from light-field data, incorporating attention mechanisms and multi-directional processing to handle inherent structural redundancies. Shin et al. [71] designed EPINet as a fully convolutional network built around epipolar geometry, establishing key architectural principles for subsequent methods. Four parallel streams process horizontal (0°), vertical (90°), and diagonal (45° , 135°) EPIs independently before fusion. This architecture learns direction-specific geometric characteristics. Each stream uses small 2×2 kernels with stride 1, appropriate for light fields’ small baselines (typical disparity range ± 4 pixels). The network achieves sub-pixel accuracy without discretizing disparity space, a fundamental limitation in traditional methods. Performance reached state-of-the-art levels with 85-fold speedup over competing optimization approaches, validating the potential of learning-based methods to achieve both superior accuracy and efficiency simultaneously.

Addressing the redundancy and variable importance in multi-view information, Du et al. developed EANet with an attention mechanism to weight different viewpoints adaptively [72]. A spatial pyramid pooling module extracts multiscale features, capturing both local detail and broader context. Three attention modes operate in tandem: per-image evaluation, mirror calculation for orthogonal directions (0° , 90°), and full four-direction analysis (0° , 45° , 90° , 135°). On the Cotton scene, performance improved 17% in MSE over EPINet ($1.64 \rightarrow 1.36$) and 14 dB in PSNR ($38.45 \rightarrow 52.60$ dB), with SSIM maintained around 0.99. These results demonstrate the value of adaptive view selection in handling redundancy and variability across different scene regions.

Chen further refined attention-based fusion with a hierarchical strategy motivated by spatially varying occlusion patterns that make uniform view aggregation suboptimal [73]. As shown in Figure 18, at the intra-branch level, channel attention weights contributions from opposite sides of the central view within each directional branch (0° , 45° , 90° , 135°). This exploits the observation that with a single occluder, occlusion affects only one side along any direction. Experimental results show that non-occluded sides receive substantially higher weights (0.869, 0.876) versus occluded sides (0.176, 0.287). This hierarchical fusion strategy

achieved top ranking on the HCI benchmark, particularly excelling with depth discontinuities and texture-poor regions, where single-level fusion approaches struggle.

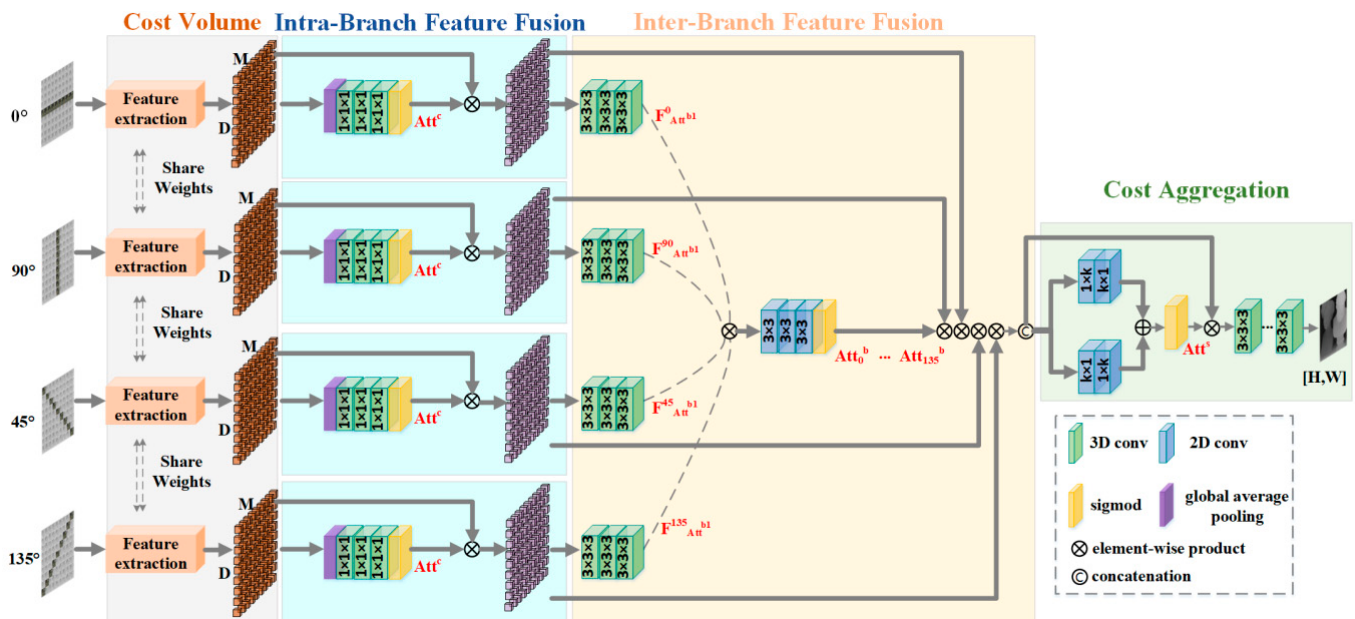


Figure 18. The architecture of the four-branch LF depth estimation network [73].

While these supervised methods demonstrate impressive performance, their reliance on ground-truth data limits applicability to real-world scenarios where accurate depth labels are difficult to obtain. Lin addressed this limitation by bridging traditional geometric constraints and deep learning through an unsupervised framework [74]. Their coarse-to-fine cascade comprises DispNet and RefineNet. The key innovation reformulates non-differentiable classical constraints, constrained angular entropy (CAE) and constrained adaptive defocus (CAD), using piecewise linear basis functions for gradient-based optimization. The adaptive angular entropy loss measures angular patch uniformity, remaining robust to occlusion through weighted entropy calculation and to noise through its statistical formulation. The adaptive defocus loss exploits the principle that incorrect disparity produces blurring in accumulated images, with blur degree correlating to estimation error. Combined with photometric consistency loss (SSIM and L1 distance), the framework matches supervised method performance on benchmark datasets while demonstrating superior generalization to unseen scenarios.

4.3. Physics-Guided Neural Networks for Dispersive Spectral Decoupling

While metasurface-based dispersive systems enable multispectral acquisition using a single sensor, conventional forward-design approaches rely on a simple phase superposition model that optimizes the phase profile only for discrete wavelengths. As a result, the axial chromatic dispersion over broad spectral ranges (e.g., 400–900 nm) exhibits strong nonlinearity ($\Delta z(\lambda) \propto \lambda^2$), preventing a linear mapping between wavelength and focal depth. This nonlinearity induces cross-talk among multispectral channels, and the coupled mechanism among axial dispersion, signal overlap, and algorithmic reconstruction remains an open problem.

1. Adjoint-based optimization for dispersion linearization: By employing adjoint optimization, the metasurface's nanostructure parameters, such as pillar height and diameter, are iteratively tuned together with the axial dispersion profile. This com-

pensates for higher-order chromatic nonlinearity, producing numerically solvable dispersive images suitable for hyperspectral reconstruction.

2. Physics-informed neural networks: Traditional data-driven deep learning models (e.g., U-Net) are often treated as “black boxes” in hyperspectral reconstruction, lacking explicit physical constraints on metasurface dispersion. This limits generalization in complex imaging scenarios. To overcome this, the axial dispersion transfer function of the metasurface, defined as the wavelength-dependent point-spread function, $PSF(\lambda, x, y)$, can be embedded within a neural network framework to construct a physics-informed inverse-convolution spectral network, achieving longitudinal-dispersion spectral decoupling.

The point-spread function (PSF) of the metalens can be calculated from Equation (1). Figure 19 shows the PSF distributions for different wavelengths from Equation (10), while Figure 20 illustrates (a) the ground-truth image, (b) the dispersive image formed by the metalens, (c) the image recovered via PSF-based deconvolution. However, the reconstructed image suffers from the loss on image details since PSFs usually work like a low-pass filter and most of high-frequency information is filtered out. In summary, while direct PSF deconvolution is physically grounded, its performance is fundamentally limited by the ill-posed nature of the inverse problem. Pure data-driven neural networks, though powerful, can lack physical consistency. The optimal approach may merge both by embedding a known PSF within a deep learning framework. This hybrid method leverages the strengths of each, using the physical model as a constraint to guide a neural network, thereby achieving superior accuracy and robustness in spectral reconstruction.

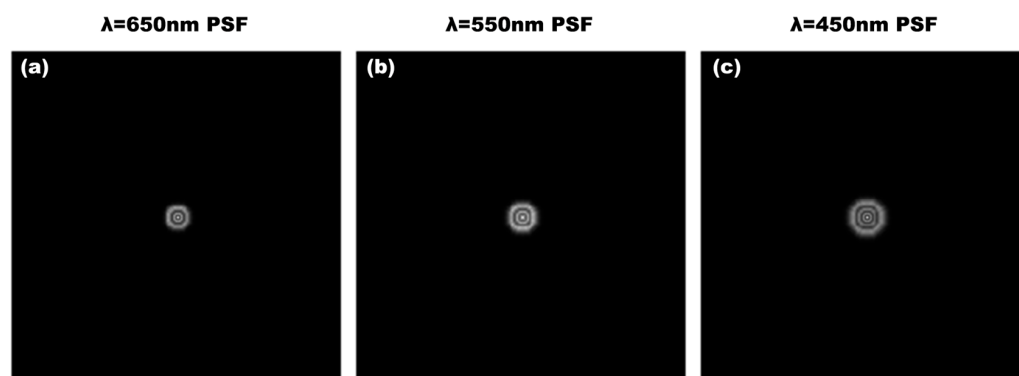


Figure 19. PSF distributions for different wavelengths. (a) PSF at $\lambda = 650$ nm. (b) PSF at $\lambda = 550$ nm. (c) PSF at $\lambda = 450$ nm.

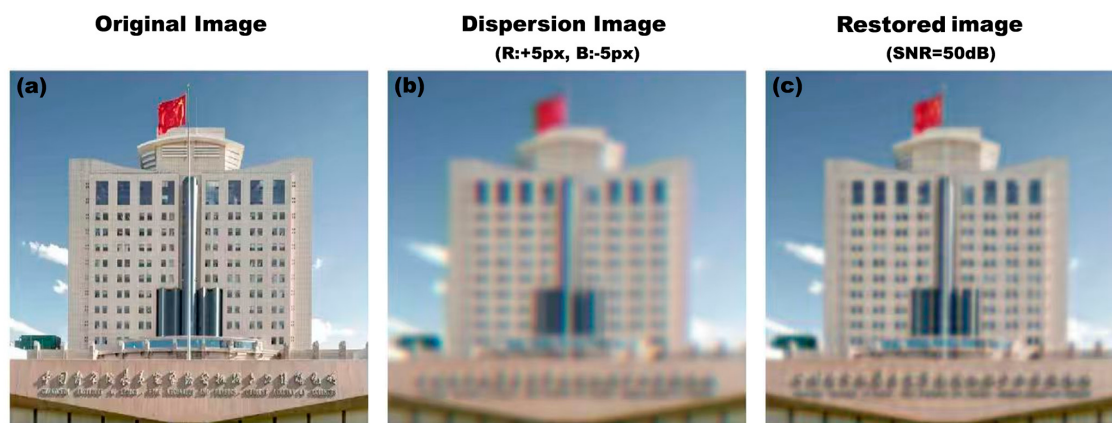


Figure 20. (a) The ground-truth image, (b) the dispersive image formed by the metalens, (c) the image recovered via PSF-based deconvolution.

4.3.1. End-to-End Co-Design Optimization Framework

Traditional design workflows for metalens imaging systems typically follow a sequential approach, where optical design precedes algorithmic compensation. However, phase discontinuities in metalenses often introduce image distortion and blurring, while the design of achromatic metalenses with large apertures and low F-numbers remains challenging. In recent years, the integration of metalenses with deep learning has opened new avenues for high-quality imaging. Among these, an end-to-end co-design framework addresses the limitations of isolated design by unifying the metalens and image restoration algorithms into a jointly optimizable system. This approach employs a differentiable geometric-phase metalens as the front end and co-optimizes it with a prior-guided image reconstruction network, enabling hardware-algorithm synergy. Experimental results demonstrate that this end-to-end joint optimization strategy significantly outperforms conventional separated design methods in terms of reconstructed image quality.

Zhang et al. [75] introduced a geometric-phase metalens into a compact snapshot hyperspectral imaging system. By co-optimizing the hardware and reconstruction algorithm, they achieved end-to-end joint training with a hyperspectral reconstruction network, significantly enhancing the overall system performance. Yu et al. [76] developed a holographic imaging system based on a metalens and computer-generated holography. By adopting an end-to-end optimization strategy to holistically optimize the diffractive optical elements and algorithms, they effectively improved imaging quality. Hu et al. [77] proposed an end-to-end jointly optimized imaging framework based on a VGG model and achromatic metalens. By simultaneously optimizing the phase profile of the metalens and the parameters of the neural network, they achieved high-quality, high-resolution image reconstruction.

The imaging system adopts an end-to-end co-design framework, which jointly optimizes the metalens and image reconstruction algorithm through differentiable forward–backward propagation, as shown in Figure 21. In the forward process, the ground-truth image is convolved with the point-spread function of the learnable metalens and corrupted by noise to generate the sensor-captured image, which is then reconstructed by a neural network. The difference between the reconstructed image and the ground-truth image constitutes the loss function. During backpropagation, the gradients of this loss function simultaneously update the neural network weights and the physical parameters of the metalens. This enables the metalens to evolve into a specialized optical encoder, while the neural network is trained concurrently as the matching decoder. Together, they cooperatively approximate the globally optimal performance of the imaging system.

The phase profile of a traditional metalens is typically defined by Equation (1), where the phase front is designed to generate a perfect spherical wavefront based on a hyperboloid phase model, with fixed coefficients that cannot be further optimized. In the end-to-end co-design approach, the phase profile of the metalens must incorporate learnable parameters, enabling the neural network to optimize these parameters during backpropagation:

$$\varphi_{lens}(x, y, z, \lambda) = -\frac{2\pi}{\lambda}(\sqrt{x^2 + y^2 + f^2} - f) + \sum_{i=0}^n a_i \left(\frac{x^2 + y^2}{R^2}\right)^i \quad (11)$$

where $\sum_{i=0}^n a_i \left(\frac{x^2 + y^2}{R^2}\right)^i$ is the learnable term, a_i represents polynomial phase factors serving as trainable parameters, and R is the radius of the metalens. The point-spread function (PSF) is defined as the intensity distribution of a point source on the image plane after passing through the metalens. It can be calculated via diffraction imaging theory as

$$PSF = \frac{1}{i\lambda z} \exp(ikz) \iint \exp(i\varphi_{lens}(x_0, y_0)) \exp\left(\frac{ik}{2z}[(x - x_0)^2 + (y - y_0)^2]\right) dx_0 dy_0 \quad (12)$$

The image formed on the sensor is a shift-invariant convolution of the ground-truth image with the PSF, which can be expressed as

$$I = \int (I_{\lambda} * P_{\lambda}) d\lambda + n_g + n_p \quad (13)$$

where I_{λ} is the ground-truth image, P_{λ} is the PSF at a certain wavelength, and n_g and n_p represent Poisson noise and Gaussian noise, respectively.

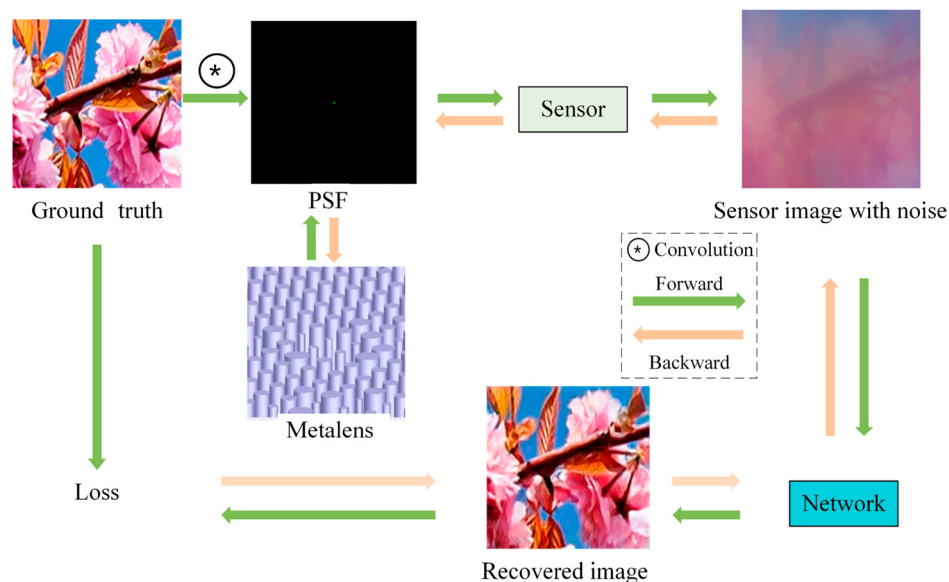


Figure 21. The end-to-end co-design optimization framework imaging reconstruction system [77].

The loss function is defined to measure the discrepancy between the restored image and the ground truth. During backpropagation, to achieve high-quality iterative image reconstruction, the parameters of the metalens phase factors and the image restoration network are continuously optimized in each iteration of the network, aiming to minimize the deviation between the restored image and the ground truth. The expression can be written as

$$\{P_{meta}, P_{NN}\} = \underset{}{argmin} \sum_{i=1}^N L(I_R, I_{\lambda}) \quad (14)$$

where P_{meta} represents the phase factor parameters of the metalens, P_{NN} denotes the parameters of the image restoration network, $L(I_R, I_{\lambda})$ is the defined loss function, I_R is the restored image, and I_{λ} is the ground-truth image.

4.3.2. Physics-Informed Attention and Generative Learning

While end-to-end frameworks optimize the pipeline, the specific architecture of the reconstruction network plays a crucial role in decoding dispersive signals. Recent research has moved beyond standard Convolutional Neural Networks (CNNs) to architectures that explicitly model long-range spectral dependencies and physical generation processes.

CNNs often struggle with the global correlation required to “unmix” signals in highly dispersive systems. Transformers, utilizing self-attention mechanisms, have emerged as a superior alternative. Cai et al. introduced MST++, a spectral-wise multi-stage Transformer that maps nonlinear dispersion relationships across channels using global attention, achieving state-of-the-art fidelity [78]. Similarly, Hu et al. proposed HDNet, which fuses high-frequency spatial features with spectral attention to prevent the smoothing of fine texture details often seen in metasurface reconstruction [79].

To tackle the ill-posed nature of 3D hyperspectral recovery, generative models are being employed to hallucinate plausible high-frequency details constrained by physical measurements. Chung et al. established the framework of Diffusion Posterior Sampling, which conditions the reverse diffusion process on the physical measurement operator (e.g., the metasurface transfer function), allowing for the rigorous solution of noisy inverse problems without task-specific training [80]. Finally, the representation of light fields is evolving from voxel grids to Implicit Neural Representations (INRs). Attal et al. demonstrated HyperReel, a 6-DoF video representation that uses ray-conditioned sampling to reconstruct high-fidelity volumetric data, offering a pathway for representing the continuous 5D light fields captured by metasurface arrays without discretization errors [81].

4.4. Hyperspectral Reconstruction Algorithm

A neural network model is proposed to incorporate the metasurface's broadband spectral response as a prior [75]. A differentiable function maps the spectral-to-spatial relationship of the PSF across the sensor plane. The differentiable reconstruction pipeline consists of three consecutive stages as metasurface phase encoding, PSF convolution simulation, and sensor-noise modeling. This framework allows the network to learn effective features and exploit the spectral power distribution for physics-based deconvolution, improving generalization across complex scenarios.

As shown in Figure 22, the feature-propagation deconvolution network achieves qualitatively superior reconstruction compared to previous methods. By combining a dispersion-optimized metalens, a PSF-informed neural deconvolution model, and a fully differentiable imaging pipeline, an end-to-end metasurface hyperspectral camera can be realized. Figure 23 illustrates the backbone of the multistage Convolutional Neural Network and an eight-layer example network, integrating PSF priors. Figure 24 shows PSFs and experimental images obtained at different stages during spectral reconstruction with physics-informed neural network. The spectral channel is 10 nm.

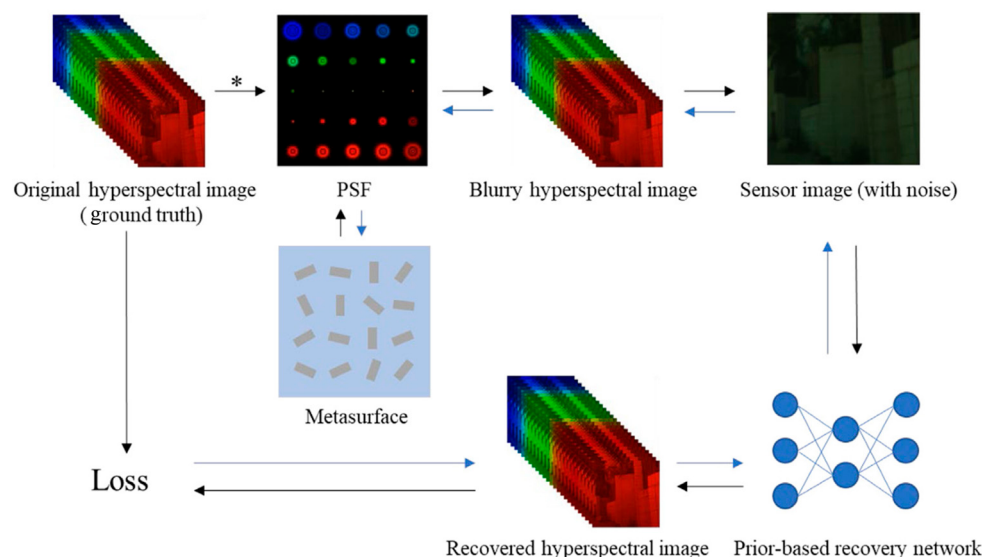


Figure 22. Optical deconvolution processing workflow (* is defined as the convolution operator and bidirectional arrows indicate a mutual interaction) [75].

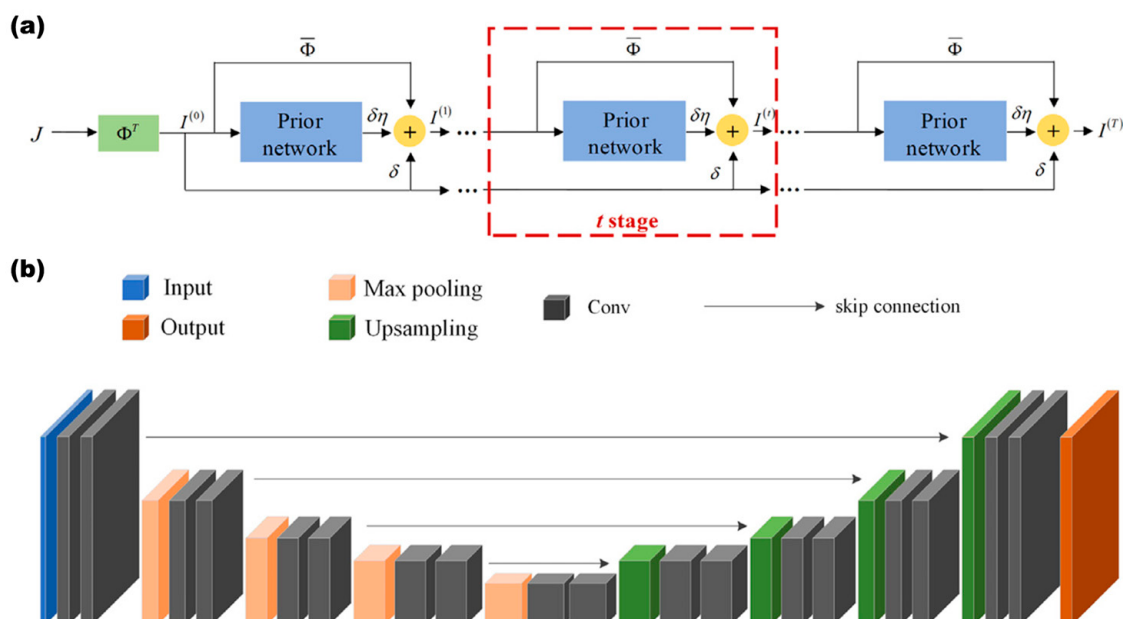


Figure 23. Neural network architecture. (a) Overall deep-learning pipeline, (b) PSF-informed prior network.

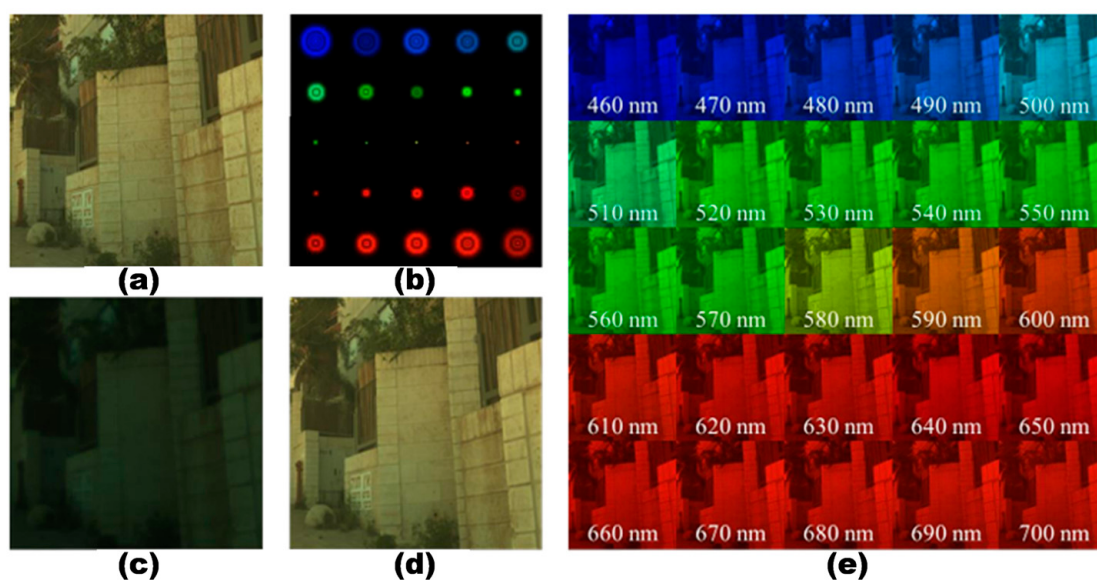


Figure 24. Images and PSFs obtained at different stages during spectral reconstruction with the physics-informed neural network. (a) The original hyperspectral image. (b) Simulated PSFs of 25 channels in the wavelength of 460~700 nm. The channel sequence is the same in (e). (c) The color image received by the imager. (d) The reconstructed hyperspectral image. (e) Reconstructed images for 25 spectral channels of 460~700 nm.

In summary, incorporating PSFs that encode multiwavelength spectral features into an inverse-convolution neural network enables robust mapping between hyperspectral characteristics and spatial images. Leveraging the precise longitudinal-dispersion control of the metasurface, this approach establishes an efficient hyperspectral reconstruction framework. Critically, achieving mathematically differentiable phase distributions across the target spectral range while maintaining optical focusing with controlled axial dispersion is essential for high-resolution spectral imaging. Efficient inverse design of such metalenses is therefore a key enabler of accurate hyperspectral reconstruction.

5. Roadmap to Implement 3D–Hyperspectral Metalens-Array Light-Field Camera

The conventional approach of using a single PSF with neural networks becomes inadequate for metasurface-based light-field cameras aiming for joint 3D and hyperspectral imaging, as spatial, angular, and spectral information are intrinsically entangled. Instead, the most promising solution involves developing a differentiable end-to-end imaging model that explicitly encapsulates the complete light transport through the metasurface array. This physics-informed framework can be effectively integrated with deep learning, either by training networks to invert this comprehensive model or by employing implicit neural representations that jointly reconstruct continuous 3D geometry and spectral radiance. By learning to invert the highly aliased sensor measurements under explicit physical constraints, this approach enables coherent reconstruction of both spatial and spectral information from a single captured image.

5.1. Coupling Mechanisms Between Metasurface Dispersion and Light-Field Imaging

Dispersion compensation in metasurfaces hinges on sophisticated nanostructural designs that match or counteract the inherent nonlinear dispersion of meta-atoms. The specific implementation mechanisms are diverse: the asymptotic phase compensation architecture (e.g., high-aspect-ratio nanopillar libraries) employs a “nonlinear matches nonlinear” strategy, optimizing the designed phase dispersion to asymptotically approach the structural dispersion, thereby enabling ultrabroadband linear control [36]. Folded-path/resonator-like configurations exploit localized interference effects within a unit cell to independently manipulate the group delay of different polarizations or wavelengths by controlling the equivalent optical path length [82]. Multi-resonance architectures synthesize a targeted quadratic phase response over a broad spectrum by superimposing multiple electromagnetic modes, achieving potent dispersion compensation [83].

However, these physical mechanisms primarily provide a “tunable parameter space.” Inverse design elevates these compensation schemes toward their performance limits by algorithmically mapping “dispersion targets” directly to “structural geometry.” It incorporates group delay ($\partial\phi/\partial\omega$), group delay dispersion ($\partial^2\phi/\partial^2\omega$), and efficiency (η) into a unified objective function, employing methods like the adjoint method or deep learning to search for the optimal nanostructure shapes and spatial distributions in a single optimization loop. This approach furnishes a deterministic physical basis for pixel-level, channel-precise wavelength encoding and separation, which is fundamental to high-performance hyperspectral imaging.

A multiphysics model that characterizes the spatial-frequency modulation of the metasurface under varying wavelengths and incident angles is therefore essential. It serves as the critical bridge connecting the incident light field to the output point-spread function (PSF). Two key challenges in this modeling are (i) dimensional extension, from the conventional 4D light field (x, y, θ, ϕ) to a 5D representation that includes the spectral dimension λ , (ii) coupling mechanism characterization, revealing how metasurface dispersion interacts with light-field propagation, such as wavelength–angle cross-modulation induced by group-velocity dispersion.

The system transfer function is obtained by superposing the electric fields of individual metasurface unit cells:

$$\mathcal{H}(u, v, \lambda, \theta) = \sum_{k=1}^K \sum_{l=1}^L [T_{kl}(\lambda, \theta) + \sum_{m=1}^M \sum_{n=1}^N C_{kl,mn}(\lambda, \theta)] \cdot e^{i\frac{2\pi}{\lambda}(u \cdot x_{kl} + v \cdot y_{kl})} \quad (15)$$

where (x_{kl}, y_{kl}) denote the center coordinates of the (k, l) -th meta-atom, $T_{kl}(\lambda, \theta)$ is the transmission with respect to (λ, θ) , and $C_{kl,mn}(\lambda, \theta)$ is the coupling between metasurface

units. (u, v) are the spatial-frequency components related to incident angles, and $n(\lambda)$ is the wavelength-dependent refractive index:

$$u = n(\lambda) \frac{\sin \theta_x}{\lambda}, \quad v = n(\lambda) \frac{\sin \theta_y}{\lambda} \quad (16)$$

After modulation by the metasurface, the object field spectrum $\mathcal{O}(u, v)$ produces the image-plane light field,

$$E_{\text{image}}(x, y, \lambda, \theta) = \mathcal{F}^{-1}[\mathcal{H}(u, v, \lambda, \theta) \cdot \mathcal{O}(u, v)] \quad (17)$$

and the corresponding point-spread function (PSF) is defined as the intensity distribution:

$$\mathcal{P}(x, y, \lambda, \theta) = |E_{\text{image}}(x, y, \lambda, \theta)|^2 \quad (18)$$

To implement the proposed roadmap, accurate characterization of the system's point-spread function (PSF)—defined as the wavelength- and angle-dependent intensity response $\mathcal{P}(x, y, \lambda, \theta)$ —is critical. We identify two complementary approaches to acquire this 5D PSF.

Physics-Based Numerical Simulation: This approach builds upon the system transfer function derived in Equation (16). The process begins with full-wave electromagnetic simulation (e.g., FDTD or RCWA) of the meta-atom library to map geometric parameters to complex transmission coefficients (S-parameters). By superposing the response of individual unit cells according to the phase profile $\varphi_{\text{lens}}(x, y, z, \lambda)$ (Equation (12)), the modulated wavefront is propagated to the sensor plane using vector diffraction theory. The simulated PSF is then calculated as in Equation (19). Iterating this process across the target spectral (VIS-NIR) and angular ranges generates the massive 5D PSF dataset required for training deep reconstruction networks without physical scanning.

Experimental Calibration: While simulation provides a noise-free baseline, experimental calibration is indispensable for capturing manufacturing imperfections (e.g., phase errors due to etching tolerances). The calibration setup typically employs a collimated tunable laser source to approximate a point source at infinity. By mechanically scanning the incident angle θ , the sensor captures the empirical PSF, $\mathcal{P}_{\text{exp}}$.

To bridge the gap between the ideal simulation and physical reality, we propose a correction framework where the system PSF is modeled as

$$\mathcal{P}_{\text{system}} = \mathcal{P}_{\text{sim}} \otimes K(\lambda, \theta) + \Delta(\lambda, \theta) \quad (19)$$

where K represents a convolutional kernel accounting for optical blur or misalignment, and Δ represents additive residual discrepancies. This hybrid approach ensures the forward model is both physically grounded and experimentally accurate.

With a comprehensive understanding of the multiphysics coupling in metalens arrays, this model allows for rapid optimization of the field distribution using inverse-design algorithms combined with adjoint-field methods, enabling efficient convergence toward the globally optimized metasurface layout.

5.2. Multidimensional Light-Field–Spectral Coupling–Decoupling and Reconstruction Theory

The end-to-end reconstruction of three-dimensional hyperspectral images from a metasurface lens array fundamentally constitutes a joint optical encoding and computational decoding information channel optimization problem. While end-to-end neural networks are widely adopted for efficient image reconstruction, conventional architectures often overlook the inherent physical constraints of the optical system, leading to decoupling errors between the PSF and the reconstructed image. Constructing a differentiable information

channel model constrained by optical encoding priors is therefore critical for establishing a neural network that integrates spectral, angular, and spatial information. This theoretical framework addresses the ill-posed nature of reconstruction in multi-physical coupled fields by enabling decoupling and recovery of spectral–angular–spatial information, thereby mitigating under-sampling and computational complexity in high-dimensional data.

(1) 5D Light-Field Tensor Representation

We define a 5D light-field tensor:

$$\mathcal{L}[i, j, m, n, p] = E_{out}(x_i, y_j, \lambda_m, \theta_n, \phi_p) \quad (20)$$

This tensor comprehensively encodes the coupled relationships across spatial (x, y), spectral (λ), and angular (θ, ϕ) domains.

(2) Core Theoretical Breakthrough: Tensor Low-Rank Decomposition

Observed data $\mathcal{Y} \in \mathbb{R}^{X \times Y \times \Lambda \times \Theta}$ are decomposed using Tucker decomposition:

$$\mathcal{Y} = \mathcal{G} \times_1 \mathbf{U}^{(X)} \times_2 \mathbf{U}^{(Y)} \times_3 \mathbf{U}^{(\Lambda)} \times_4 \mathbf{U}^{(\Theta)} + \mathcal{E} \quad (21)$$

where $\mathcal{G}$ is the core tensor capturing intermodal coupling strength, $\mathbf{U}^{(*)}$ are factor matrices describing basis functions along each dimension, and the decomposition ranks ($R_X, R_Y, R_\Lambda, R_\Theta$) are determined via cross-validation.

(3) Joint Optimization Reconstruction Model

A spectral–angular–spatial joint objective function is formulated as

$$\min \|\mathcal{Y} - \mathcal{P} * \mathcal{L}\|_F^2 + \alpha \|\nabla_\lambda \mathcal{S}\|_1 + \beta \text{TV}_{xy}(\mathcal{L}) + \gamma \text{TV}_{\theta\phi}(\mathcal{L}) \quad (22)$$

where $\mathcal{S} = \mathcal{F}_\lambda \mathcal{L}$ is a spectrally sparse transform and the total variation (TV) regularization suppresses spatial–angular noise:

$$\text{TV}_{xy}(\mathcal{L}) = \sum_{i,j} \sqrt{|\nabla_x \mathcal{L}_{i,j}|^2 + |\nabla_y \mathcal{L}_{i,j}|^2} \quad (23)$$

$$\text{TV}_{\theta\phi}(\mathcal{L}) = \sum_{i,j} \sqrt{|\nabla_\theta \mathcal{L}_{i,j}|^2 + |\nabla_\phi \mathcal{L}_{i,j}|^2} \quad (24)$$

(4) Wavelength–Angular Decoupling

Introducing a decoupling operator D , the coupled tensor is decomposed as

$$\mathcal{L} = D(\mathcal{L}_\lambda \otimes \mathcal{L}_\theta) + \mathcal{N} \quad (25)$$

where $\mathcal{L}_\lambda \in \mathbb{R}^{X \times Y \times \Lambda}$ encodes pure spectral features, $\mathcal{L}_\theta \in \mathbb{R}^{X \times Y \times \Theta}$ encodes pure angular features, and the decoupling error $\mathcal{N}$ is <5% (quantified by mutual information).

(5) End-to-End Neural Network Implementation

The feasibility of end-to-end reconstruction lies in effectively coupling the physical dispersion characteristics of the metasurface lens with the data-driven feature learning of neural networks. Traditional stepwise reconstruction methods artificially separate optical degradation from algorithmic recovery, resulting in cumulative errors, particularly under high-spectral-mixing or low-SNR conditions [84]. The end-to-end framework embeds the dispersion transfer function of the metasurface lens directly into differentiable neural network layers, forcing the model to simultaneously learn optical degradation rules and reconstruction priors during training. This strategy overcomes the intrinsic limitations of conventional workflows. While this paradigm is promising, its success hinges on the accuracy of the differentiable forward model and the availability of large-scale training datasets. Future work must focus on bridging the “sim-to-real” gap through sophisticated

physical modeling introduced above and efficient self-supervision strategies that minimize the need for experimentally captured ground-truth 5D data.

5.3. Implementation Pathways and Challenges for Seed Phenomics

This section will review existing computational algorithms for light-field and spectral decoupling, and on this basis, discuss the potential pathways and challenges for their co-design with metasurface array hardware. This integration forms the technical core of the future platform we propose.

5.3.1. Key Technology Map, Challenges, and Future Perspectives

To bridge the gap between proof-of-concept devices and a practical 5D phenomics platform, we first define a “Target Design Envelope” based on the intrinsic biological requirements of seed screening. For example, detecting micro-cracks in maize or soybean seeds (typically 5–15 mm in diameter) mandates a spatial resolution of approximately 30–50 μm [1,2]. Simultaneously, accurate chemical phenotyping (e.g., protein and moisture content) requires extending the spectral range into the NIR-SWIR region (up to 1700 nm) [6]. Even though research on the application of metalenses or metalens arrays in plant and seed phenomics is still in its early stages, with few studies reported, we anticipate this field will experience rapid growth, driving significant advancements and even revolutionizing traditional imaging methodologies [85].

Table 1 highlights the advantages, challenges and trends of metasurface optics with computational imaging as hyperspectral imaging, 3D reconstruction and fusion applications. Key challenges include optical efficiency of metasurfaces, spectral cross-talk, ill-posed reconstruction from mixed signals, computational scalability, standardization of multimodal phenotyping data, and robustness in non-laboratory environments. Future research is expected to leverage advanced materials, dedicated hardware accelerators (e.g., GPUs), and AI-driven pipeline optimization to address these challenges, enabling scalable, high-throughput seed phenotyping platforms.

Table 1. Fundamental metasurface technology for 3D reconstruction and hyperspectral imaging.

Technology Direction	Main Advantages	Technical Challenges	Future Trends
Metasurfaces for Hyperspectral Imaging	• Single-shot acquisition: Complete spectral data without mechanical scanning	• Broadband efficiency: Maintaining high transmission across visible–NIR spectrum	• New dielectric materials for enhanced efficiency
	• System miniaturization: Replaces complex spectroscopic systems	• Spectral cross-talk: Interwavelength interference elimination	• Physics-informed neural networks
	• Controllable dispersion: Precise wavelength manipulation via nanostructures	• Manufacturing precision: Nanoscale fabrication tolerances affecting spectral response	• Nanoimprinting for mass production
	• High design freedom: Inverse design for optimal spectral encoding	• Calibration complexity: Accurate PSF wavelength dependence characterization	• Integrated optoelectronic computing
	• Computational imaging fusion: Co-optimization with deep learning algorithms	• Computational demand: High-dimensional data reconstruction requirements	• Transition from lab to industrial applications
Metasurface Arrays for 3D Reconstruction	• Multi-view synchronization: Simultaneous capture from different perspectives	• Viewpoint consistency: Performance uniformity across array elements	• Large-scale uniform array fabrication
	• Depth information extraction: Accurate 3D reconstruction via parallax calculation	• Spatial resolution limitation: Resolution limited by microlens array size	• Advanced depth estimation algorithms
	• Motion capture: Dynamic scene capture correlated in single shot	• Depth accuracy: Reconstruction precision constrained by baseline and resolution	• Wide-field metasurface design
	• Occlusion recovery: Occlusion potentially reconstructed from angular data	• Data reconstruction: Complex algorithm for image alignment and reconstruction	• Real-time processing architectures

Table 1. Cont.

Technology Direction	Main Advantages	Technical Challenges	Future Trends
Metasurface Arrays for 3D-Hyperspectral Fusion	• 5D information acquisition: Simultaneous spatial (x, y), angular (θ, φ), spectral (λ) capture	• 5D coupling decoupling: Extreme complexity in spatial-spectral-angular separation	• Multiphysics-informed machine learning
	• Hardware-algorithm co-design: End-to-end joint optimization	• System modeling difficulty: Accurate multiphysics coupling models	• Advanced tensor decomposition methods
	• Complementary information enhancement: Mutual constraints improve reconstruction	• Data explosion: Massive 5D data storage, transmission, and processing	• Heterogeneous integration technologies
	• Single-exposure completeness: Full high-dimensional data cube without scanning	• Manufacturing limits: Large-scale high-uniformity metasurface array fabrication	• Edge computing for real-time applications
	• Compact multifunctionality: Multiple traditional systems in single device	• Computational requirements: Real-time processing needs powerful hardware	• Next-generation computational photography

A central challenge in realizing this roadmap is the “Sim-to-Real” gap. Training end-to-end networks requires massive datasets of paired measurements (sensor images) and ground truth data (5D spectral-spatial volumes). However, acquiring “Ground-Truth 5D Data” for biological samples like seeds is physically prohibitive; no existing instrument can simultaneously capture high-resolution 3D geometry and hyperspectral data in a single snapshot to serve as a label.

To overcome this, we propose a two-stage data strategy:

Physics-Driven Synthetic Data Generation: Utilizing the rigorous forward model described in Section 5.1, we can generate synthetic sensor measurements from existing databases (e.g., 3D scans of seeds combined with USDA spectral libraries). This allows the network to learn the fundamental physics of dispersion and light-field encoding.

Domain Adaptation: To generalize to real-world conditions, transfer learning techniques should be applied. The network pre-trained on simulation data can be fine-tuned using a small set of “unpaired” real data or “weakly paired” data (e.g., separate 3D and spectral measurements of the same seed variety), employing cycle-consistent generative adversarial networks (CycleGANs) to align the domain distributions.

To illustrate the advantages of metasurface-based approaches over traditional methods, Table 2 provides a detailed comparison of key metrics, including volume/thickness, chromatic aberration control, system integration, and imaging efficiencies, across refractive microlens arrays, metasurface array, CASSI systems, and metasurface modulators.

Table 2. Comparative analysis of key technologies in snapshot hyperspectral imaging: refractive microlens arrays, metasurface arrays, coded aperture snapshot spectral imaging (CASSI), and on-chip metasurface modulators.

Category	Metric (Reviewer Requested)	Refractive MLA [86,87]	Meta-Array [88,89]	CASSI [90,91]	Metasurface [22,92]
1. Dispersion Mechanisms	Axial Dispersion Sensitivity ($\partial t/\partial \lambda$)	~0.005 $\mu\text{m}/\text{nm}$	Negligible (Achromatic/Corrected)	N/A	Non-Uniformly Coupled
	Lateral Dispersion Rate ($\Delta x/\Delta \lambda$)	~0.005 pixels/nm	N/A (Spatial Multiplexing)	~0.5–2 pixels/nm	Low/Parasitic (Radial Blur if Uncorrected)
2. Optics and Efficiency	Chromatic Aberration Control Range	Broadband (430–780 nm)	Discrete Bands (Specific Center λ)	Broadband (430–780 nm Material Limited)	Broadband (400–1700 nm)
	Monochromatic Efficiency	Very High (>95%)	High (~60–80%)	High (~80–90%)	High (~70–90%)
	Broadband Imaging Efficiency	High (>90%)	Medium (~40–60%)	Medium (~50%, Mask Blocking)	High (~70%)
	Throughput (NA)	0.2–0.5	Medium (NA~0.3)	Low–Medium (NA Limited by F-Number)	High (NA up to 0.8–0.9)
3. Physical Specs	Thickness/Volume	Bulky (TTL > 50 mm)	Ultra-Compact (TTL < 500 μm)	Bulky System (TTL > 100 mm)	Ultra-Compact (TTL < 1 mm)
	System Integration (Component Count)	Low (Lens Group + MLA + Sensor)	High (CMOS Integration)	Low (Obj + Mask + Relay + Prism + Sensor)	High (Monolithic Integration)

Table 2. Cont.

Category	Metric (Reviewer Requested)	Refractive MLA [86,87]	Meta-Array [88,89]	CASSI [90,91]	Metasurface [22,92]
4. Imaging Specs	Spectral Resolution	Low (~20–50 nm)	Medium (~10–20 nm)	Medium–High (<10 nm)	High (~1–5 nm)
	Spatial Resolution	Low (Trade-off: Spatial/Spectral)	Medium (Sub-Sampled)	Medium (Mask Resolution Limited)	High (Full Sensor Pixel Count)
	Depth Res./DOF	High/Refocusable (Light-Field Capability)	Low (Fixed Focus)	Low/Fixed (Planar Imaging)	High/Tunable (Point-Spread Function Engineering)
	Signal-to-Noise Ratio (SNR)	High (>40 dB)	Medium (~30 dB)	Medium (20–30 dB, Shot Noise)	Medium (25–35 dB)
	Calibration Burden	Geometric (Microlens Center Alignment)	Light (Filter Response Matrix)	Heavy (Full 3D PSF Scanning)	Light (1D z-λ Curve Fitting)
5. Algorithm and Data	Reconstruction Runtime	Fast (Linear Processing, <5 s)	Real-time (Demosaicing, ms)	Slow (Iterative/DL, mins to hours)	Real-Time (~100 ms)
	Dataset/Training Req.	None	Low (For Demosaicing)	High (Model-Based/DL-Based Reconstruction)	High (96 Channels, Training Augmentation/Transfer Learning)

5.3.2. Validation Framework for Seed Phenotyping

To ensure the proposed metasurface platform yields reliable biological insights, we establish a structured validation protocol specific to seed traits. This protocol assesses performance across three hierarchical levels.

Level 1: Data Fidelity (Ground Truthing):

The reconstructed point clouds will be compared against high-precision Micro-CT scans (resolution < 10 μm). Accuracy will be quantified using the Chamfer Distance (CD) and volumetric error percentages. Reconstructed spectra at distinct spatial points will be validated against measurements from a calibrated contact spectrometer. Metrics include Root Mean Square Error (RMSE) and Spectral Angle Mapper (SAM) to assess spectral shape fidelity.

Level 2: Trait Extraction Accuracy:

Morphological traits (e.g., length, width, surface area) extracted from the metasurface system will be regressed against manual caliper or machine-vision benchmark measurements.

Level 3: Functional Prediction (Reproducibility and Application):

The ultimate validation lies in the platform’s predictive power. We propose training regression models (e.g., PLS-R) on the reconstructed data to predict biochemical traits (e.g., protein content, moisture). These predictions will be cross-validated against destructive laboratory chemical analysis (e.g., Kjeldahl method for protein). System robustness will be evaluated via the Intraclass Correlation Coefficient (ICC) by imaging the same seed batches under varying orientations and lighting conditions.

Overall, this roadmap outlines a pathway toward a transformative, metasurface-driven imaging platform for next-generation seed phenomics research. Its privilege advantage lies in leveraging a single, compact optical front-end to replace complex multi-component systems, enabling the non-destructive, single-shot acquisition of fused morphological and biochemical data. This integrated approach directly addresses the critical bottleneck in high-throughput seed phenomics by providing a multidimensional “phenotypic fingerprint” that bridges the gap between genotype and observable traits.

However, the practical implementation of this vision faces several significant limitations. Optical efficiency of large-area metasurfaces, particularly across broad spectral bands, remains a challenge, constraining signal-to-noise ratio [52]. The computational burden of solving the ill-posed inverse problem for 5D data reconstruction is immense, demanding further innovation in algorithms and hardware acceleration [93]. While industrial sorting demands high throughput, most broadband metasurfaces currently exhibit optical efficiencies of 40–60% (vs. >90% for refractive optics) [52]. To compensate for this photon loss, higher illumination intensity is often required to maintain the signal-to-noise

ratio (SNR). This creates a conflict with the thermal limit ($\sim 40^\circ\text{C}$) of viable seeds, as excessive heat accumulation from high-intensity light can damage seed germinability [94]. Therefore, improving broadband efficiency is not merely an optical goal but a biological necessity to enable low-dose, high-speed imaging. Unlike discrete refractive filters, dispersive metasurfaces often exhibit spectral–spatial cross-talk. This complexity increases the calibration burden and can degrade the quantitative accuracy of chemical inversion (e.g., oil/protein estimation) [95]. Consequently, the roadmap emphasizes the need for advanced deep-learning reconstruction algorithms to decouple these signals effectively.

Future developments will likely focus on overcoming these hurdles through advanced inverse design incorporating novel optical materials and structure to boost efficiency, and the creation of “physics-informed” models that deeply integrate optical principle with optimization algorithm for faster and more robust reconstruction. Ultimately, the convergence of metasurface optics, computational imaging is poised to create intelligent, all-in-one imagers that will not only revolutionize seed phenotyping but also find broad applications in biomedical diagnostics, remote sensing, and consumer electronics [96].

6. Conclusions and Perspectives

This review has articulated the potential applications of seed phenotyping, with metasurface-based light-field platform for 3D reconstruction and hyperspectral imaging. We began by elucidating the fundamental principles of both technologies: light-field imaging, which captures the angular distribution of light to enable computational depth estimation and 3D seed model generation, and hyperspectral imaging, which resolves the spectral signature of a scene to map its molecular composition. The convergence of these two modalities promises a powerful, non-destructive tool for high-throughput seed phenomics, capable of capturing both morphological and physiological traits in a single snapshot. Our discussion focused on the pivotal role of metasurfaces as the promising technology for this integration. Key design technologies were explored, including advanced inverse design and adjoint optimization methods, which allow for the creation of metalens arrays tailored for specific light-field and spectral encoding tasks.

Regarding the current state of the art, we surveyed representative implementations, such as lateral-dispersion systems for compact spectral imaging and axial-dispersion designs that encode depth and wavelength into a single measurement. The critical challenge of computational reconstruction was addressed, highlighting the emergence of physics-informed neural networks that embed the system’s point-spread function as a prior to robustly decouple the intertwined spatial, angular, and spectral information from a single encoded image.

Future development of this field is poised to follow a clear roadmap. The initial focus on establishing high-fidelity 3D imaging with achromatic metalens arrays will progressively integrate spectral capabilities, ultimately achieving a fully fused 3D–hyperspectral data volume. While challenges in optical efficiency, computational complexity, and system robustness remain, the ongoing synergy between metasurface design, computational algorithms, and machine learning is set to unlock compact, high-throughput, and multidimensional imaging platforms. These systems will not only advance seed phenomics but also broadly impact precision agriculture, biomedical imaging, and beyond, marking a significant leap forward in non-destructive analytical technologies.

Author Contributions: Conceptualization, J.Y. and Q.S.; methodology, Q.Z.; validation, S.L.; investigation, J.G.; resources, F.G.; data curation, S.W.; writing—original draft preparation, J.Y.; writing—review and editing, Q.S.; visualization, Q.H.; supervision, Q.L.; project administration, C.L.; funding acquisition, M.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by [Yuelushan Laboratory Breeding Program] grant number [YLS-2025-ZY03022], and [Changsha Municipal Key Special Projects] grant number [kq2404013].

Data Availability Statement: Data will be made available on request.

Conflicts of Interest: Author Qi Song was employed by the company Soleilware Photonics LLC. Author Chao Li was employed by the company Qualcomm. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Shiferaw, B.; Prasanna, B.M.; Hellin, J.; Bänziger, M. Crops that feed the world 6. Past successes and future challenges to the role played by maize in global food security. *Food Secur.* **2011**, *3*, 307–327. [\[CrossRef\]](#)
- Erenstein, O.; Jaleta, M.; Sonder, K.; Mottaleb, K.; Prasanna, B.M. Global maize production, consumption and trade: Trends and R&D implications. *Food Secur.* **2022**, *14*, 1295–1319. [\[CrossRef\]](#)
- Jin, B.; Qi, H.; Jia, L.; Tang, Q.; Gao, L.; Li, Z.; Zhao, G. Determination of viability and vigor of naturally-aged rice seeds using hyperspectral imaging with machine learning. *Infrared Phys. Technol.* **2022**, *122*, 104097. [\[CrossRef\]](#)
- Qi, H.; Huang, Z.; Sun, Z.; Tang, Q.; Zhao, G.; Zhu, X.; Zhang, C. Rice seed vigor detection based on near-infrared hyperspectral imaging and deep transfer learning. *Front. Plant Sci.* **2023**, *14*, 1283921. [\[CrossRef\]](#)
- Wang, X.; Zhu, M.; Li, J.; Yang, Y.; Xie, H.; Duan, Y.; Cao, N.; Kan, R.; Yu, Y. Evaluation and Development Trends of Optical Detection Technology for Seed Vigor. *Spectroscopy* **2025**. [\[CrossRef\]](#)
- Gao, T.; Chandran, A.K.N.; Paul, P.; Walia, H.; Yu, H. HyperSeed: An End-to-End Method to Process Hyperspectral Images of Seeds. *Sensors* **2021**, *21*, 8184. [\[CrossRef\]](#) [\[PubMed\]](#)
- Yang, W.; Feng, H.; Zhang, X.; Zhang, J.; Doonan, J.H.; Batchelor, W.D.; Xiong, L.; Yan, J. Crop Phenomics and High-Throughput Phenotyping: Past Decades, Current Challenges, and Future Perspectives. *Mol. Plant* **2020**, *13*, 187–214. [\[CrossRef\]](#)
- ElMasry, G.; Mandour, N.; Al-Rejaie, S.; Belin, E.; Rousseau, D. Recent Applications of Multispectral Imaging in Seed Phenotyping and Quality Monitoring—An Overview. *Sensors* **2019**, *19*, 1090. [\[CrossRef\]](#)
- Klasen, D.; Fischbach, A.; Sydoruk, V.; Kochs, J.; Bühler, J.; Koller, R.; Huber, G. Seed-to-plant-tracking: Automated phenotyping of seeds and corresponding plants of Arabidopsis. *Front. Plant Sci.* **2025**, *16*, 1539424. [\[CrossRef\]](#)
- Zhang, M.; Song, J.; Jia, H.; Zhang, X.; Yang, W.; Wang, Y.; Wang, H. Prediction of Vigor of Naturally Aged Seeds from Xishuangbanna Cucumber (*Cucumis sativus* L. var. *xishuangbannanensis*) Using Hyperspectral Imaging. *Agriculture* **2025**, *15*, 1043. [\[CrossRef\]](#)
- Danilevicz, M.F.; Bayer, P.E.; Nestor, B.J.; Bennamoun, M.; Edwards, D. Resources for image-based high-throughput phenotyping in crops and data sharing challenges. *Plant Physiol.* **2021**, *187*, 699–715. [\[CrossRef\]](#)
- Berry, J.C.; Fahlgren, N.; Pokorny, A.A.; Bart, R.S.; Veley, K.M. An automated, high-throughput method for standardizing image color profiles to improve image-based plant phenotyping. *PeerJ* **2018**, *6*, e5727. [\[CrossRef\]](#)
- Reuzeau, C.; Pen, J.; Frankard, V.; de Wolf, J.; Peerbolte, R.; Broekaert, W.; van Camp, W. TraitMill: A Discovery Engine for Identifying Yield-enhancement Genes in Cereals. *Plant Gene Trait.* **2010**, *1*, 1–6. [\[CrossRef\]](#)
- Guo, Q.; Wu, F.; Pang, S.; Zhao, X.; Chen, L.; Liu, J.; Xue, B.; Xu, G.; Li, L.; Jing, H.; et al. Crop 3D—A LiDAR based platform for 3D high-throughput crop phenotyping. *Sci. China Life Sci.* **2018**, *61*, 328–339. [\[CrossRef\]](#)
- Wu, D.; Guo, Z.; Ye, J.; Feng, H.; Liu, J.; Chen, G.; Zheng, J.; Yan, D.; Yang, X.; Xiong, X.; et al. Combining high-throughput micro-CT-RGB phenotyping and genome-wide association study to dissect the genetic architecture of tiller growth in rice. *J. Exp. Bot.* **2019**, *70*, 545–561. [\[CrossRef\]](#) [\[PubMed\]](#)
- Qian, Y.; Cao, P.; Yin, W.; Dai, F.; Hu, F.; Yan, Z. Calculation method of surface shape feature of rice seed based on point cloud. *Comput. Electron. Agric.* **2017**, *142*, 416–423. [\[CrossRef\]](#)
- Yang, J.; Zhao, Q.; Song, Q.; Zhu, M.; Liu, S.; Yu, Y.; Wang, H. Co-optimized tunable-focus light field imaging system for 3D seed phenotyping: From optical design to computational reconstruction. *Optica Open* **2025**, preprint.. [\[CrossRef\]](#)
- Xiong, Z.; Liu, S.; Tan, J.; Huang, Z.; Li, X.; Zhuang, G.; Fang, Z.; Chen, T.; Zhang, L. Combining Hyperspectral Techniques and Genome-Wide Association Studies to Predict Peanut Seed Vigor and Explore Associated Genetic Loci. *Int. J. Mol. Sci.* **2024**, *25*, 8414. [\[CrossRef\]](#)
- Yan, H.; Zhang, Z.; Lv, Y.; Nie, Y. Integrated multispectral imaging, germination phenotype, and transcriptomic analysis provide insights into seed vigor responsive mechanisms in quinoa under artificial accelerated aging. *Front. Plant Sci.* **2024**, *15*, 1435154. [\[CrossRef\]](#)
- Li, Y.; Li, P.; Zheng, X.; Liu, H.; Zhao, Y.; Sun, X.; Liu, W.; Zhou, S. Design of a Novel Microlens Array and Imaging System for Light Fields. *Micromachines* **2024**, *15*, 1166. [\[CrossRef\]](#) [\[PubMed\]](#)

21. Pan, W.; Umana-Membreno, G.A.; Akhavan, N.D.; Tan, H.H.; Neshev, D.; Wesemann, L.; Leslie, P.; Driggers, R.; Faraone, L. Design and Simulation of Metalens Arrays for Enhanced MWIR Imaging Array Performance. *J. Electron. Mater.* **2025**, *54*, 8304–8314. [\[CrossRef\]](#)
22. Fan, Z.-B.; Qiu, H.-Y.; Zhang, H.-L.; Pang, X.-N.; Zhou, L.-D.; Liu, L.; Ren, H.; Wang, Q.-H.; Dong, J.-W. A broadband achromatic metalens array for integral imaging in the visible. *Light Sci. Appl.* **2019**, *8*, 67. [\[CrossRef\]](#)
23. World's First Hyper-Spectral Color Imaging by a Metalens Camera Without Chromatic Aberration. Available online: <https://group.ntt/en/newsrelease/2022/10/24/221024a.html> (accessed on 24 October 2022).
24. Dhanya, V.G.; Subeesh, A.; Susmita, C.; Amaresh; Saji, S.J.; Dilsha, C.; Keerthi, C.; Nunavath, A.; Singh, A.N.; Kumar, S. High throughput phenotyping using hyperspectral imaging for seed quality assurance coupled with machine learning methods: Principles and way forward. *Plant Physiol. Rep.* **2024**, *29*, 749–768. [\[CrossRef\]](#)
25. Wang, R.-F.; Qu, H.-R.; Su, W.-H. From sensors to insights: Technological trends in image-based high-throughput plant phenotyping. *Smart Agric. Technol.* **2025**, *12*, 101257. [\[CrossRef\]](#)
26. Hu, X.; Xu, W.; Fan, Q.; Yue, T.; Yan, F.; Lu, Y.; Xu, T. Metasurface-based computational imaging: A review. *Adv. Photonics* **2024**, *6*, 014002. [\[CrossRef\]](#)
27. Hu, X.; Li, Z.; Miao, L.; Fang, F.; Jiang, Z.; Zhang, X. Measurement Technologies of Light Field Camera: An Overview. *Sensors* **2023**, *23*, 6812. [\[CrossRef\]](#)
28. Lin, B.; Tian, Y.; Zhang, Y.; Zhu, Z.; Wang, D. Deep learning methods for high-resolution microscale light field image reconstruction: A survey. *Front. Bioeng. Biotechnol.* **2024**, *12*, 1500270. [\[CrossRef\]](#)
29. Ghamisi, P.; Yokoya, N.; Li, J.; Liao, W.; Liu, S.; Plaza, J.; Rasti, B.; Plaza, A. Advances in Hyperspectral Image and Signal Processing: A Comprehensive Overview of the State of the Art. *IEEE Geosci. Remote Sens. Mag.* **2018**, *5*, 37–78. [\[CrossRef\]](#)
30. Liu, Y.; Pu, H.; Sun, D.-W. Hyperspectral imaging technique for evaluating food quality and safety during various processes: A review of recent applications. *Trends Food Sci. Technol.* **2017**, *69*, 25–35. [\[CrossRef\]](#)
31. Levoy, M.; Ng, R.; Adams, A.; Footer, M.; Horowitz, M. Light Field Microscopy. In *Acm Siggraph 2006 Papers*; ACM: New York, NY, USA, 2006; pp. 924–934.
32. Broxton, M.; Grosenick, L.; Yang, S.; Cohen, N.; Andalman, A.; Deisseroth, K.; Levoy, M. Wave optics theory and 3-D deconvolution for the light field microscope. *Opt. Express* **2013**, *21*, 25418–25439. [\[CrossRef\]](#)
33. Zhang, Z.; Bai, L.; Cong, L.; Yu, P.; Zhang, T.; Shi, W.; Li, F.; Du, J.; Wang, K. Imaging volumetric dynamics at high speed in mouse and zebrafish brain with confocal light field microscopy. *Nat. Biotechnol.* **2021**, *39*, 74–83. [\[CrossRef\]](#)
34. Khorasaninejad, M.; Capasso, F. Metalenses: Versatile multifunctional photonic components. *Science* **2017**, *358*, eaam8100. [\[CrossRef\]](#)
35. Berry, M.V. Quantal phase factors accompanying adiabatic changes. *Proc. R. Soc. London. Ser. A Math. Phys. Sci.* **1984**, *392*, 45–57. [\[CrossRef\]](#)
36. Hu, Y.; Jiang, Y.; Zhang, Y.; Yang, X.; Ou, X.; Li, L.; Kong, X.; Liu, X.; Qiu, C.-W.; Duan, H. Asymptotic dispersion engineering for ultra-broadband meta-optics. *Nat. Commun.* **2023**, *14*, 6649. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Li, S.-H.; Sun, C.; Tang, P.-Y.; Liao, J.-H.; Hsieh, Y.-H.; Fung, B.-H.; Fang, Y.-H.; Kuo, W.-H.; Wu, M.-H.; Chang, H.-C.; et al. Augmented reality system based on the integration of polarization-independent metalens and micro-LEDs. *Opt. Express* **2024**, *32*, 11463–11473. [\[CrossRef\]](#)
38. Lin, R.J.; Su, V.-C.; Wang, S.; Chen, M.K.; Chung, T.L.; Chen, Y.H.; Kuo, H.Y.; Chen, J.-W.; Chen, J.; Huang, Y.-T.; et al. Achromatic metalens array for full-colour light-field imaging. *Nat. Nanotechnol.* **2019**, *14*, 227–231. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Fan, Q.; Xu, W.; Hu, X.; Zhu, W.; Yue, T.; Zhang, C.; Yan, F.; Chen, L.; Lezec, H.J.; Lu, Y.; et al. Trilobite-inspired neural nanophotonic light-field camera with extreme depth-of-field. *Nat. Commun.* **2022**, *13*, 2130. [\[CrossRef\]](#) [\[PubMed\]](#)
40. Zaidi, A.; Rubin, N.A.; Meretska, M.L.; Li, L.W.; Dorrah, A.H.; Park, J.-S.; Capasso, F. Metasurface-enabled single-shot and complete Mueller matrix imaging. *Nat. Photonics* **2024**, *18*, 704–712. [\[CrossRef\]](#)
41. Zhao, Z.; Liu, X.; Ji, Y.; Zhang, Y.; Chen, Y.; Luo, Z.; Song, Y.; Geng, Z.; Tanaka, T.; Qi, F.; et al. Meta-lens digital image correlation. *Opto-Electronic Adv.* **2025**, *8*, 250014-1–250014-12. [\[CrossRef\]](#)
42. Song, Y.; Yuan, J.; Chen, Q.; Liu, X.; Zhou, Y.; Cheng, J.; Xiao, S.; Chen, M.K.; Geng, Z. Three-dimensional varifocal meta-device for augmented reality display. *Photonix* **2025**, *6*, 6. [\[CrossRef\]](#)
43. Zhou, Y.; Chen, Z.; Cheng, J.; Zhang, Q.; Geng, Z.; Wu, Z.; Chen, M.K. High-Resolution 3D Imaging with Tunable Point Cloud Projection Based on Meta-Device. *Laser Photonics Rev.* **2025**, e01327. [\[CrossRef\]](#)
44. Fröch, J.E.; Colburn, S.; Brady, D.J.; Heide, F.; Veeraraghavan, A.; Majumdar, A. Computational imaging with meta-optics. *Optica* **2025**, *12*, 774. [\[CrossRef\]](#)
45. Wang, P.; Mohammad, N.; Menon, R. Chromatic-aberration-corrected diffractive lenses for ultra-broadband focusing. *Sci. Rep.* **2016**, *6*, 21545. [\[CrossRef\]](#)
46. Yu, H.; Xie, Z.; Li, C.; Li, C.; de SMenezes, L.; Maier, S.A.; Ren, H. Dispersion Engineering of Metalenses. *Appl. Phys. Lett.* **2023**, *123*, 240503. [\[CrossRef\]](#)

47. He, Y.; Chen, M.K.; Huang, M.; Zhang, Y.; Liu, X.; Luo, Z.; Yao, C.; Li, H.; Zeng, F.; Geng, Z.; et al. Dispersive Meta-lens Thermometry for High-temperature Measurements. *Nat. Commun.* **2025**, *16*, 10090. [\[CrossRef\]](#)
48. Tan, S.; Yang, F.; Boominathan, V.; Veeraraghavan, A.; Naik, G.V. 3D Imaging Using Extreme Dispersion in Optical Metasurfaces. *Acs Photonics* **2021**, *8*, 1421–1429. [\[CrossRef\]](#)
49. Liu, Y.; Li, W.-D.; Xin, K.-Y.; Chen, Z.-M.; Chen, Z.-Y.; Chen, R.; Chen, X.-D.; Zhao, F.-L.; Zheng, W.-S.; Dong, J.-W. Ultra-wide FOV meta-camera with transformer-neural-network color imaging methodology. *Adv. Photonics* **2024**, *6*, 056001. [\[CrossRef\]](#)
50. Hua, X.; Wang, Y.; Wang, S.; Zou, X.; Zhou, Y.; Li, L.; Yan, F.; Cao, X.; Xiao, S.; Tsai, D.P.; et al. Ultra-compact snapshot spectral light-field imaging. *Nat. Commun.* **2022**, *13*, 2732. [\[CrossRef\]](#) [\[PubMed\]](#)
51. Roques-Carnes, C.; Wang, K.; Yang, Y.; Majumdar, A.; Lin, Z. Computational Metaoptics for Imaging. *arXiv* **2024**. [\[CrossRef\]](#)
52. Chen, W.T.; Park, J.-S.; Marchioni, J.; Millay, S.; Yousef, K.M.A.; Capasso, F. Dispersion-engineered metasurfaces reaching broadband 90% relative diffraction efficiency. *Nat. Commun.* **2023**, *14*, 2544. [\[CrossRef\]](#)
53. Sun, J.; Wei, K.; Eboli, T.; Wang, C.; Zheng, C.; Zhou, Z.; Majumdar, A.; Heidrich, W.; Heide, F. Collaborative On-Sensor Array Cameras. *ACM Trans. Graph.* **2025**, *44*, 55. [\[CrossRef\]](#)
54. Yang, J.; Cui, K.; Huang, Y.; Zhang, W.; Feng, X.; Liu, F. Angle-Insensitive Spectral Imaging Based on Topology-Optimized Plasmonic Metasurfaces. *Laser Photonics Rev.* **2024**, *18*, 2400255. [\[CrossRef\]](#)
55. Bao, Y.; Li, B. Single-shot simultaneous intensity, phase and polarization imaging with metasurface. *Natl. Sci. Rev.* **2025**, *12*, nwae418. [\[CrossRef\]](#)
56. Zeng, Y.; Zhong, H.; Long, Z.; Cao, H.; Jin, X. From performance to structure: A comprehensive survey of advanced metasurface design for next-generation imaging. *npj Nanophotonics* **2025**, *2*, 39. [\[CrossRef\]](#)
57. Li, Z.; Pestourie, R.; Lin, Z.; Johnson, S.G.; Capasso, F. Empowering Metasurfaces with Inverse Design: Principles and Applications. *ACS Photonics* **2022**, *9*, 2178–2192. [\[CrossRef\]](#)
58. Quan, D.; Liu, X.; Tang, Y.; Liu, H.; Min, S.; Li, G.; Srivastava, A.K.; Cheng, X. Dielectric Metalens by Multilayer Nanoimprint Lithography and Solution Phase Epitaxy. *Adv. Eng. Mater.* **2023**, *25*, 2201824. [\[CrossRef\]](#)
59. Moon, S.-W.; Kim, Y.; Yoon, G.; Rho, J. Recent Progress on Ultrathin Metalenses for Flat Optics. *iScience* **2020**, *23*, 101877. [\[CrossRef\]](#)
60. Choi, M.; Kim, J.; Moon, S.; Shin, K.; Nam, S.-W.; Park, Y.; Kang, D.; Jeon, G.; Lee, K.-I.; Yoon, D.H.; et al. Roll-to-plate printable RGB achromatic metalens for wide-field-of-view holographic near-eye displays. *Nat. Mater.* **2025**, *24*, 535–543. [\[CrossRef\]](#) [\[PubMed\]](#)
61. Pestourie, R.; Pérez-Arancibia, C.; Lin, Z.; Shin, W.; Capasso, F.; Johnson, S.G. Inverse design of large-area metasurfaces. *Opt. Express* **2018**, *26*, 33732–33747. [\[CrossRef\]](#)
62. Li, Z.; Pestourie, R.; Park, J.-S.; Huang, Y.-W.; Johnson, S.G.; Capasso, F. Inverse design enables large-scale high-performance meta-optics reshaping virtual reality. *Nat. Commun.* **2022**, *13*, 2409. [\[CrossRef\]](#)
63. Gao, R.-Q.; Song, Q.; Liu, H.; Gao, J.-B.; Wang, X.-Y.; Bayanheshig; Li, C.; Ding, G.; Gong, Y.; Chen, X.-H.; et al. Design of near-Infrared reconfigurable metalens on Silicon-On-Insulator (SOI) platform with Fabry–Perrot phase shifter. *Opt. Commun.* **2019**, *446*, 56–63. [\[CrossRef\]](#)
64. Ng, R.; Levoy, M.; Brédif, M.; Duval, G.; Horowitz, M.; Hanrahan, P. Light Field Photography with a Hand-Held Plenoptic Camera. Ph.D. Thesis, Stanford University, Stanford, CA, USA, 2005.
65. Levoy, M. Light Fields and Computational Imaging. *Computer* **2006**, *39*, 46–55. [\[CrossRef\]](#)
66. Jin, D.; Zhang, S.; Huo, X.; Zhang, W.; Yang, F. A Two-Step Calibration Method for Unfocused Light Field Camera Based on Projection Model Analysis. *arXiv* **2021**. [\[CrossRef\]](#)
67. Wanner, S.; Goldluecke, B. Variational Light Field Analysis for Disparity Estimation and Super-Resolution. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *36*, 606–619. [\[CrossRef\]](#) [\[PubMed\]](#)
68. Zhang, S.; Sheng, H.; Li, C.; Zhang, J.; Xiong, Z. Robust depth estimation for light field via spinning parallelogram operator. *Comput. Vis. Image Underst.* **2016**, *145*, 148–159. [\[CrossRef\]](#)
69. Sabater, N.; Seifi, M.; Drazic, V.; Sandri, G.; Pérez, P. Accurate Disparity Estimation for Plenoptic Images. In Proceedings of the European Conference on Computer Vision, Zurich, Switzerland, 6 September 2014; Springer International Publishing: Cham, Switzerland; pp. 548–560.
70. Wang, T.-C.; Efros, A.A.; Ramamoorthi, R. Occlusion-Aware Depth Estimation Using Light-Field Cameras. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 3487–3495.
71. Shin, C.; Jeon, H.-G.; Yoon, Y.; Kweon, I.S.; Kim, S.J. EPINET: A Fully-Convolutional Neural Network Using Epipolar Geometry for Depth from Light Field Images. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 19–21 June 2018; pp. 4748–4757.
72. Du, Y.; Zhang, Q.; Hua, D.; Hou, J.; Wang, B.; Zhu, S.; Zhang, Y.; Fang, Y. EANet: Depth Estimation Based on EPI of Light Field. *BioMed Res. Int.* **2021**, *2021*, 8293151. [\[CrossRef\]](#)
73. Chen, J.; Zhang, S.; Lin, Y. Attention-based Multi-Level Fusion Network for Light Field Depth Estimation. *Proc. AAAI Conf. Artif. Intell.* **2021**, *35*, 1009–1017. [\[CrossRef\]](#)

74. Lin, L.; Li, Q.; Gao, B.; Yan, Y.; Zhou, W.; Kuruoglu, E.E. Unsupervised learning of light field depth estimation with spatial and angular consistencies. *Neurocomputing* **2022**, *501*, 113–122. [\[CrossRef\]](#)
75. Zhang, Q.; Yu, Z.; Liu, X.; Wang, C.; Zheng, Z. End-to-end joint optimization of metasurface and image processing for compact snapshot hyperspectral imaging. *Opt. Commun.* **2023**, *530*, 129154. [\[CrossRef\]](#)
76. Yu, Z.; Zhang, Q.; Tao, X.; Li, Y.; Tao, C.; Wu, F.; Wang, C.; Zheng, Z. High-performance full-color imaging system based on end-to-end joint optimization of computer-generated holography and metalens. *Opt. Express* **2022**, *30*, 40871–40883. [\[CrossRef\]](#)
77. Hu, S.; Shi, R.; Wang, B.; Wei, Y.; Qi, B.; Zhou, P. Full-Color Imaging System Based on the Joint Integration of a Metalens and Neural Network. *Nanomaterials* **2024**, *14*, 715. [\[CrossRef\]](#)
78. Cai, Y.; Lin, J.; Lin, Z.; Wang, H.; Zhang, Y.; Pfister, H.; Timofte, R.; Van Gool, L. MST++: Multi-stage Spectral-wise Transformer for Efficient Spectral Reconstruction. In Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), New Orleans, LA, USA, 18–24 June 2022; pp. 744–754.
79. Hu, X.; Cai, Y.; Lin, J.; Wang, H.; Yuan, X.; Zhang, Y.; Timofte, R.; Van Gool, L. HDNet: High-resolution Dual-domain Learning for Spectral Compressive Imaging. In Proceedings of the 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 18–24 June 2022; pp. 17521–17530.
80. Chung, H.; Kim, J.; Mccann, M.T.; Klasky, M.L.; Ye, J.C. Diffusion Posterior Sampling for General Noisy Inverse Problems. *arXiv* **2024**. [\[CrossRef\]](#)
81. Attal, B.; Huang, J.-B.; Richardt, C.; Zollhöfer, M.; Kopf, J.; O’Toole, M.; Kim, C. HyperReel: High-Fidelity 6-DoF Video with Ray-Conditioned Sampling. In Proceedings of the 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, BC, Canada, 17–24 June 2023; pp. 16610–16620.
82. Zhang, F.; Bao, H.; Pu, M.; Guo, Y.; Kang, T.; Li, X.; He, Q.; Xu, M.; Ma, X.; Luo, X. Dispersion-engineered spin photonics based on folded-path metasurfaces. *Light Sci. Appl.* **2025**, *14*, 198. [\[CrossRef\]](#)
83. Tsilipakos, O.; Koschny, T. Multiresonant metasurfaces for arbitrarily broad bandwidth pulse chirping and dispersion compensation. *Phys. Rev. B* **2023**, *107*, 165408. [\[CrossRef\]](#)
84. Chen, Y.; Gui, X.; Zeng, J.; Zhao, X.-L.; He, W. Combining Low-Rank and Deep Plug-and-Play Priors for Snapshot Compressive Imaging. *IEEE Trans. Neural Netw. Learn. Syst.* **2023**, *35*, 16396–16408. [\[CrossRef\]](#) [\[PubMed\]](#)
85. Hu, J.; Yang, W. Metalens array miniaturized microscope for large-field-of-view imaging. *Opt. Commun.* **2023**, *555*, 130231. [\[CrossRef\]](#)
86. Nussbaum, P.; Völkel, R.; Herzig, H.P.; Eisner, M.; Haselbeck, S. Design, fabrication and testing of microlens arrays for sensors and microsystems. *Pure Appl. Opt. J. Eur. Opt. Soc. Part A* **1997**, *6*, 617–636. [\[CrossRef\]](#)
87. Dansereau, D.G.; Pizarro, O.; Williams, S.B. Decoding, Calibration and Rectification for Lenselet-Based Plenoptic Cameras. In Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Portland, OR, USA, 23–28 June 2013; pp. 1027–1034.
88. Bian, L.; Wang, Z.; Zhang, Y.; Li, L.; Zhang, Y.; Yang, C.; Fang, W.; Zhao, J.; Zhu, C.; Meng, Q.; et al. A broadband hyperspectral image sensor with high spatio-temporal resolution. *Nature* **2024**, *635*, 73–81. [\[CrossRef\]](#)
89. Horie, Y.; Arbabi, A.; Arbabi, E.; Kamali, S.M.; Faraon, A. Wide bandwidth and high resolution planar filter array based on DBR-metasurface-DBR structures. *Opt. Express* **2016**, *24*, 11677–11682. [\[CrossRef\]](#)
90. Wagadarikar, A.; John, R.; Willett, R.; Brady, D. Single disperser design for coded aperture snapshot spectral imaging. *Appl. Opt.* **2008**, *47*, B44–B51. [\[CrossRef\]](#) [\[PubMed\]](#)
91. Gehm, M.E.; John, R.; Brady, D.J.; Willett, R.M.; Schulz, T.J. Single-shot compressive spectral imaging with a dual-disperser architecture. *Opt. Express* **2007**, *15*, 14013–14027. [\[CrossRef\]](#) [\[PubMed\]](#)
92. Faraji-Dana, M.; Arbabi, E.; Arbabi, A.; Kamali, S.M.; Kwon, H.; Faraon, A. Compact folded metasurface spectrometer. *Nat. Commun.* **2018**, *9*, 4196. [\[CrossRef\]](#)
93. Han, X.-H.; Wang, J.; Jiang, H. Recent Advancements in Hyperspectral Image Reconstruction from a Compressive Measurement. *Sensors* **2025**, *25*, 3286. [\[CrossRef\]](#)
94. Bewley, J.D.; Bradford, K.; Hilhorst, H. *Seeds: Physiology of Development, Germination and Dormancy*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2012.
95. Yang, Z.; Albrow-Owen, T.; Cai, W.; Hasan, T. Miniaturization of Optical Spectrometers. *Science* **2021**, *371*, 0722. [\[CrossRef\]](#) [\[PubMed\]](#)
96. Li, Q.; Liang, M.; Liu, S.; Liu, J.; Chen, S.; Wen, S.; Luo, H. Phase reconstruction via metasurface-integrated quantum analog operation. *Opto-Electronic Adv.* **2025**, *8*, 240239–1–240239-9. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.