APPLIED PHYSICS

Miniature color camera via flat hybrid meta-optics

Samuel Pinilla[1], Johannes E. Fröch[2], Seyyed Reza Miri Rostami[3], Vladimir Katkovnik[3]*, Igor Shevkunov[3], Arka Majumdar[2], Karen Egiazarian[3]*

The race for miniature color cameras using flat meta-optics has rapidly developed the end-to-end design framework using neural networks. Although a large body of work has shown the potential of this methodology, the reported performance is still limited due to fundamental limitations coming from meta-optics, mismatch between simulated and resultant experimental point spread functions, and calibration errors. Here, we use a HIL optics design methodology to solve these limitations and demonstrate a miniature color camera via flat hybrid meta-optics (refractive + meta-mask). The resulting camera achieves high-quality full-color imaging for a 5-mm aperture optics with a focal length of 5 mm. We observed a superior quality of the images captured by the hybrid meta-optical camera compared to a compound multi-lens optics of a mirrorless commercial camera.

INTRODUCTION

Modern cameras are complex systems optimized for high-quality imaging, typically composed of multiple lenses to overcome chromatic and geometric aberrations. This complexity, however, often comes at the expense of size and weight (*1*). Contrarily, low-weight, ultra-thin, and small cameras are urgently needed for integration within mobile platforms, such as next-generation smartphones, drones, or point-of-care medical devices (*2*). Thus, the miniaturization of cameras, while maintaining high image quality, has become a major driving force in optics and photonics research. A solution to this conundrum is computational imaging, where a digital backend augments the deficiencies of the optical components and improves the image quality, which has thus become a multidisciplinary research field at the intersection of optics, mathematics, and digital image processing (*3*, *4*).

In parallel, progress over recent years has also advanced the field of flat diffractive optical elements (DOE). While DOEs have traditionally been used for non-imaging applications, several recent works have demonstrated their capabilities in imaging. Especially meta-optics, a class of subwavelength DOEs, are well-suited for computational imaging, as they allow to impart arbitrary freeform phase-profiles (*5*, *6*) on the incident wavefront with high spatial resolution (*7*–*9*, *11*, *12*). In conjunction with the right computational backend, meta-optics can produce high-quality images, as recently shown (*2*), albeit with a small aperture (500 μm). An especially promising approach is to co-optimize the meta-optics and the computational backend in an "end-to-end" design framework, where the hardware and software are equally considered, thus ensuring optimal system-level performance (*7*, *10*, *13*, *14*).

There are three potential advantages of meta-optics in computational imaging: (i) they can achieve important size and weight reduction (thickness in micrometres) (*15*); (ii) performance beyond conventional systems is achievable in extended imaging modalities, such as hyperspectral imaging, extended-depth-of-field (*16*–*18*), and face recognition (*22*); (iii) nearly arbitrary subwavelength resolution manipulation of wavefields is possible due to the immense progress in nanofabrication over the recent decades. Despite these promises, the image quality of captured images using large-aperture (> 1mm) meta-optics is plagued with strong chromatic and geometric aberrations. It has been shown that there are fundamental limitations in capturing broadband images solely using meta-optics (*1*, *19*). While complicated meta-atom engineering can help, fabricating large aperture meta-optics with small features remains challenging. Last, the design tools to create large aperture meta-optics are extremely computationally expensive. This poses a serious challenge, as the current state-of-the-art end-to-end design approach (*10*, *20*, *21*) largely involves numerical modeling of the optics.

In this work, we circumvent all these challenges by using a hybrid refractive/ meta-optics system in tandem with a computational backend, following a recently proposed design methodology (*3*). The desired phase profile is first optimized by a hardware-in-the-loop (HIL) strategy, where the DOE, implemented by a spatial light modulator (SLM), is configured and updated. At the same time, a refractive lens and a sensor remain fixed. This not only bypasses the expensive computational requirements to design the optics but also directly accounts for sensor noise and precludes calibration errors. The optimized phase mask is then implemented using meta-optics (details in Materials and Methods). Using hybrid optics over meta-optics–only systems improves the low light efficiency of meta-optics (*23*), aberrations, and reconstruction quality (*24*, *25*). We note that we use simple meta-atoms to ensure fabricability, and the meta-atoms are designed only at 510 nm, where the SLM operates. Although only a single wavelength is considered, we still achieve high-quality full-color imaging, as the meta-optics extend the depth of focus (*26*). Thus, even using simple meta-atoms optimized for only one wavelength, we can realize high-quality broadband imaging thanks to hybrid optics and a computational backend. The resulting imaging systems have a focal length and aperture of 5 mm. To illustrate the potential of this camera, we compare it with a single lens-only configuration and the compound multi-lens optics of the Sony Alpha 1 III mirrorless commercial camera. This comparison confirms the efficiency of the meta-hybrid optics in full-color imaging, whereas our designed hybrid

[1]Science and Technology Facilities Council, Harwell, UK. [2]University of Washington, Seattle, WA, USA. [3]Tampere University, Tampere, Finland.
*Corresponding author. Email: vladimir.katkovnik@tuni.fi (V.K.); karen.eguiazarian@tuni.fi (K.E.)

meta-optics camera volume is approximately 108× smaller than the mirrorless commercial setup.

RESULTS

Designed hybrid meta-optics

The design of full-color, large aperture hybrid optics is prohibitively expensive, due to the extensive number of scatters involved in the meta-optics. We bridge this massive gap from theoretical modeling to real-life application using a HIL configuration, summarized in Fig. 1A, where we use an end-to-end design framework to optimize the phase distribution of the DOE and the corresponding computational backend. Here, the lenses L_1 and L_2 form the 4f-telescopic system projecting wavefront from the SLM plane to the imaging lens L_3, complementary metal-oxide semiconductor is a registering camera. d_1 is the distance between the scene and the plane of the hybrid optics, and d_2 is a distance between the optics and the sensor. Thus, the wavefront modulated by SLM is projected directly to the lens plane, what exactly corresponds to the hybrid optical setup, where the phase mask is attached to the lens. Thus, we ensure that there is no air gap between meta-mask (SLM) and lens.

The pixel-wise programmable phase SLM in Fig. 1A acts as a placeholder to implement the phase distribution of the DOE, which is iteratively updated, while a computational backend is co-optimized. Fundamentally, the HIL methodology guarantees proper "modeling" of the camera black-box, including DOE, refractive optic, and sensor, as the discrepancies between model and physical reality are eliminated. Essentially, the physical experiment in HIL replaces the approximate mathematical models of wavefront propagation through the diffractive optical hybrid. Further details on the optimization can be found in Materials and Methods and the Supplementary Materials. This optimization yields a phase distribution as shown in Fig. 1B following the methodology presented in (*27*). Although captured images in this configuration are blurred, the co-optimized computational backend recovers the image and yields an all-in-focus full-color image with high quality. This is directly demonstrated in the same configuration on the example of a three-dimensional (3D) scene, whereas features at various depths are magnified to emphasize the high imaging quality.

We then translated the phase distribution into a meta-optics consisting of nano-pillars, which locally impart the desired phase delay, with further details on design and fabrication given in Materials and Methods. Although we only considered a single wavelength (511 nm) in the design and implementation, we show full-color operation in this hybrid configuration. The finalized meta-optics (diameter of 5 mm) and its properties are summarized in Fig. 2A, where scanning electron micrographs (bottom row) show the structural integrity of the nano-pillars across the meta-optic. Further, to compare and validate the phase response of the meta-optic, we used holographic characterization, following (*28*). The measured phase across the entire device shows uniformity among sections, and a line cut of the profile (middle) directly indicates a good match between design and fabricated device. The wavelength used for this characterization is 511 nm. We note that the discrepancy between the design and fabrication is inevitable as there will always be fabrication imperfection. However, thanks to the computational backend, we can achieve a fairly robust solution, which does not rely on ultra-sophisticated fabrication techniques.

Optical properties of the image formation and elements of the camera (meta-mask, lens, sensor, and fixed data-processing

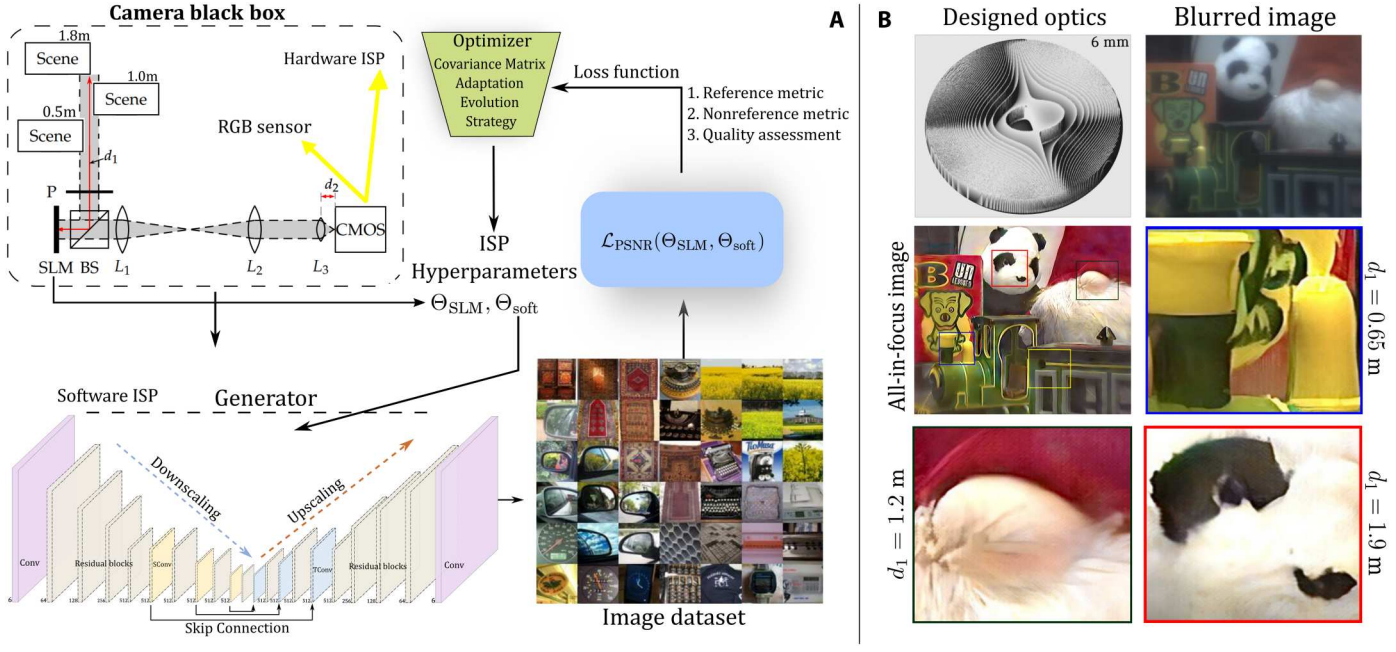

Fig. 1. Schematic hardware-in-the-loop methodology to co-optimize hybrid meta-optics and inverse imaging in an end-to-end framework. (**A**) The camera black-box model is composed of a SLM with phase modulation updated in optimization iterations and fixed sensor and hardware ISP. The output of this black box forms the input for the image reconstruction algorithm. (**B**) SLM phase distribution after optimization (for 10-mm focal distance) and imaging results for an all-in-focus scenario. The captured image is blurred, a crisp image is obtained after reconstruction as highlighted by details at various distances from the sensor. CMOS, complementary metal-oxide semiconductor.

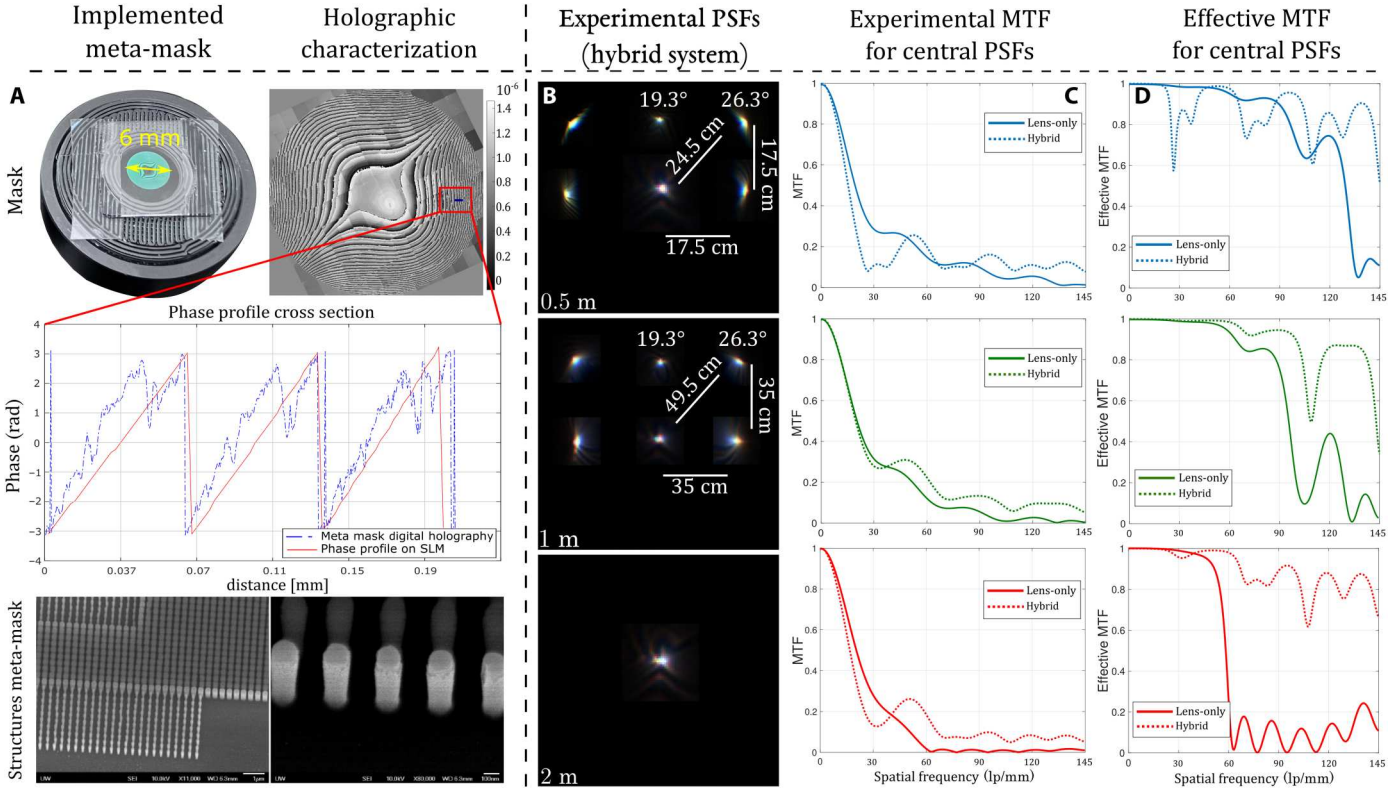

Fig. 2. Holographic characterization of the meta-mask phase profile along with its PSFs across depths and MTFs per channels. (**A**) Properties of the fabricated meta-optics showing a photograph of the meta-mask on a 3D-printed mount. The bottom row show scanning electron micrographs highlighting the structural integrity across the device. To determine the quality of the implemented phase, we perform a holographic characterization of the meta-optic, indicating a close match between design (red curve) and measurement (blue). (**B**) Experimental PSFs of the hybrid system for distances d_1 = 0.5,1.0, and 2.0 m. (**C** and **D**) Experimental and effective MTFs (EMTFs) for central PSFs of hybrid and lens-only systems.

software) are characterized by the point spread functions (PSFs). These PSFs are found experimentally using a broadband laser as a point-wise light source. For this purpose, we used the flashlight of the smartphone along with an optical fiber (Quartz, 200 μm/0.22 numerical aperture, SMA to SMA) as a point source with an almost uniform response in the spectral range (400 to 700) nm in a dark place. The projections of this light beam on different points of the optical system as the sensor register them are shown in Fig. 2B. These intensity PSFs are given for distances d_1 = 0.5,1.0, and 2.0 m (from top to bottom). The hybrid system exhibits PSFs, which are quite similar for various depths, thus enabling the extended depth-of-focus functionality provided by the meta-mask.

In Fig. 2 (C and D), we also report the modulation transfer functions (MTFs) and the effective MTFs (EMTFs) for the lens-only and hybrid systems per RGB channels. By definition, the MTFs are calculated as the modulus of the optical transfer functions (OTFs), which are the Fourier transform of the measured PSFs. Furthermore, the EMTFs are computed here as products of MTFs and the modulus of the Wiener inverse of the OTFs. In this way, the EMTFs take into account the whole optical system plus imaging postprocessing. We use these Wiener inverse-based EMTFs to demonstrate the potential of the designed meta-optics hybrid to characterize imaging quality. The calculations of MTFs and EMTFs are shown for RGB data, for the focal distance d_1 = 1.0 m, and for the central point of the sensor.

The principal advantage of hybrid optics versus lens-only is clear. Note that the pass-band of the hybrid optics is higher than that for the camera with the lens-only systems. The EMTFs for the hybrid take values closer to 1, thus enabling higher imaging quality for the hybrid system. In particular, the lens-only system for the red-spectral MTF takes much lower and even zero values. Note also that the spectral EMTFs are well consolidated, i.e., they are close to each other for all spatial frequencies for the RGB spectrum. These results suggest that the inverse imaging for a hybrid system is achromatic with strongly suppressed chromatic aberrations.

Imaging results

To assess the performance of the hybrid refractive meta-optics, we captured images in two configurations: (1) three monitor imaging with three fixed distances between the monitors and sensor (Fig. 3) and a (2) real-life scene with arbitrary locations of various colored objects at several distances relative to the camera (Fig. 4). First, we compare the imaging performance of the hybrid meta-optics with a single-lens camera in configuration 1, where an image (ground truth) is displayed at a monitor at distances of either 0.5, 1, or 1.8 m, relative to the sensor. As shown in the inset of Fig. 3, the meta-optics was placed directly in front of the refractive lens to achieve the hybrid configuration, whereas for reference measurements (lens-only), the meta-optics was removed. In all cases, the images are

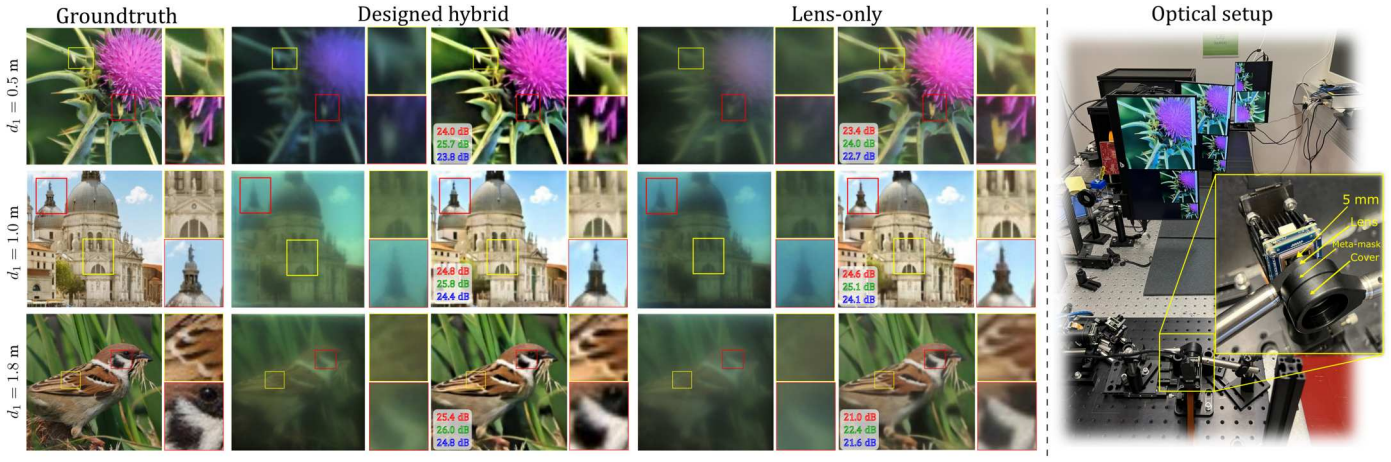

Fig. 3. Three monitor imaging setup using the designed hybrid meta-optics The first column shows examples of the displayed ground truth at distances of d_1 = 0.5,1.0, and 1.8 m from top to bottom row, respectively. Specific details of the images are highlighted and shown as zoomed-in sections next to the image. Columns 2 and 3 show the imaging results obtained with the hybrid and lens-only systems, with raw captures and computationally reconstructed images shown left and right, respectively. The PSNR values are reported for each depth and color channel separately and shown as inset in the reconstructed images. The optical setup used for image capture is shown on the very right of the figure. The inset shows an enlarged view of the hybrid camera, consisting of the refractive lens placed at the focal length of 5 mm in front of the sensor and the meta-optics placed directly in front of the refractive lens.

displayed before and after computational reconstruction. As visually evident by comparison of specific image details, the hybrid optics yields better images, specifically for distances of 0.5 and 1.8 m. The superior performance of the hybrid is further quantitatively underlined by the calculated peak signal-to-noise ratio (PSNR) values, which are higher across all color channels and all distances.

In configuration 2, we further compare the hybrid optics with a compound multi-lens camera, namely, a Sony Alpha 1 III sensor with a Sony SEL85F18 lens (at the aperture limit values of *F*/1.8 and *F*/22). The captured images (for a fixed focus distance of 1 m) and reconstructed images are shown in rows 2 and 3, for the hybrid and the single lens camera, in columns 3 and 4, respectively. Specific image details, for objects at out-of-focus distances of 0.6 and 1.1 m, are enlarged in rows 4 and 5. As in the prior case, images obtained with a hybrid optics exhibit sharper images with a better resolution of details at both depths. Further, in comparison with the compound camera and aperture of *F*/1.8, the hybrid shows clearly better imaging quality across a larger depth range. Even for the smaller aperture value of *F*/22, the images obtained with the hybrid exhibit sharper details for out-of-focus depths at 0.6 m.

DISCUSSION

In this work, a full-color miniature computational camera, composed of a refractive lens and large aperture meta-optics, is demonstrated. The hybrid solution has an a low *F*-number of 1.0 obtained by the ratio $\frac{5\text{ mm}}{5\text{ mm}} = \frac{\text{focaldistance}}{\text{aperturediameter}}$. A crucial aspect in the development is the design of the meta-optic, where we used a HIL methodology to optimize the phase distribution and the computational backend in an end-to-end design framework. Specifically, an SLM acts as a placeholder for the meta-optics to iteratively optimize the phase distribution and the corresponding computational backend to achieve full-color high-quality imaging over a large training set. The only target metric considered is the quality of the deconvolved RGB image, which deviates from existing methods that optimize a digital differentiable model of the camera and the computational back-end. Thus, we entirely bypass the requirement to optimize the meta-optics phase distribution with an electromagnetic solver, which otherwise poses a computationally prohibitively expensive problem due to the large number of scatterer involved. Also, this methodology directly accounts for the fixed properties of the refractive lens and the sensor, therefore yielding a result corresponding to reality and a more robust computational backend.

It is essential to point out that the computational cost of state-of-the-art design methodologies increases rapidly as the device dimensions increase (*9*). As mentioned in (*9*), the most expensive computational step in terms of time and memory to design optics using state-of-the-art methods is the computation of the PSFs of the camera. Therefore, since our HIL methodology is a PSF-less design strategy because the light propagation is not numerically modeled but is physically done, we are decreasing the computational cost by at least two orders of magnitude. The time needed to process physical observations is about 1.01 s (meaning real-time design, out of reach by state-of-the-art methodologies). In addition, traditional end-to-end algorithms to design optics have a memory consumption of approximately 20 GB (*9*), which we do not need to use in our HIL method since PSFs are not computed. Therefore, we are decreasing the memory consumption by at least one order of magnitude, measured in gigabytes. The memory consumption of our design methodology is about 3 GB.

The optimized phase distribution is then directly translated to a meta-optics using simple square pillars in SiN (700 nm in height, 350-nm period) on a quartz substrate (500-μm thickness), which does not add a substantial thickness to the refractive lens (4.5-mm thickness). Although we consider the phase response for only a single wavelength (511 nm), the resultant meta-optics still yields an improvement over the entire visible range as validated in full-color imaging experiments using a monitor setup (Fig. 3), as well as under ambient light (Fig. 4). In comparison to a single refractive lens-only configuration, we achieve better image quality, particularly for off-focus depths. Even compared with a commercial

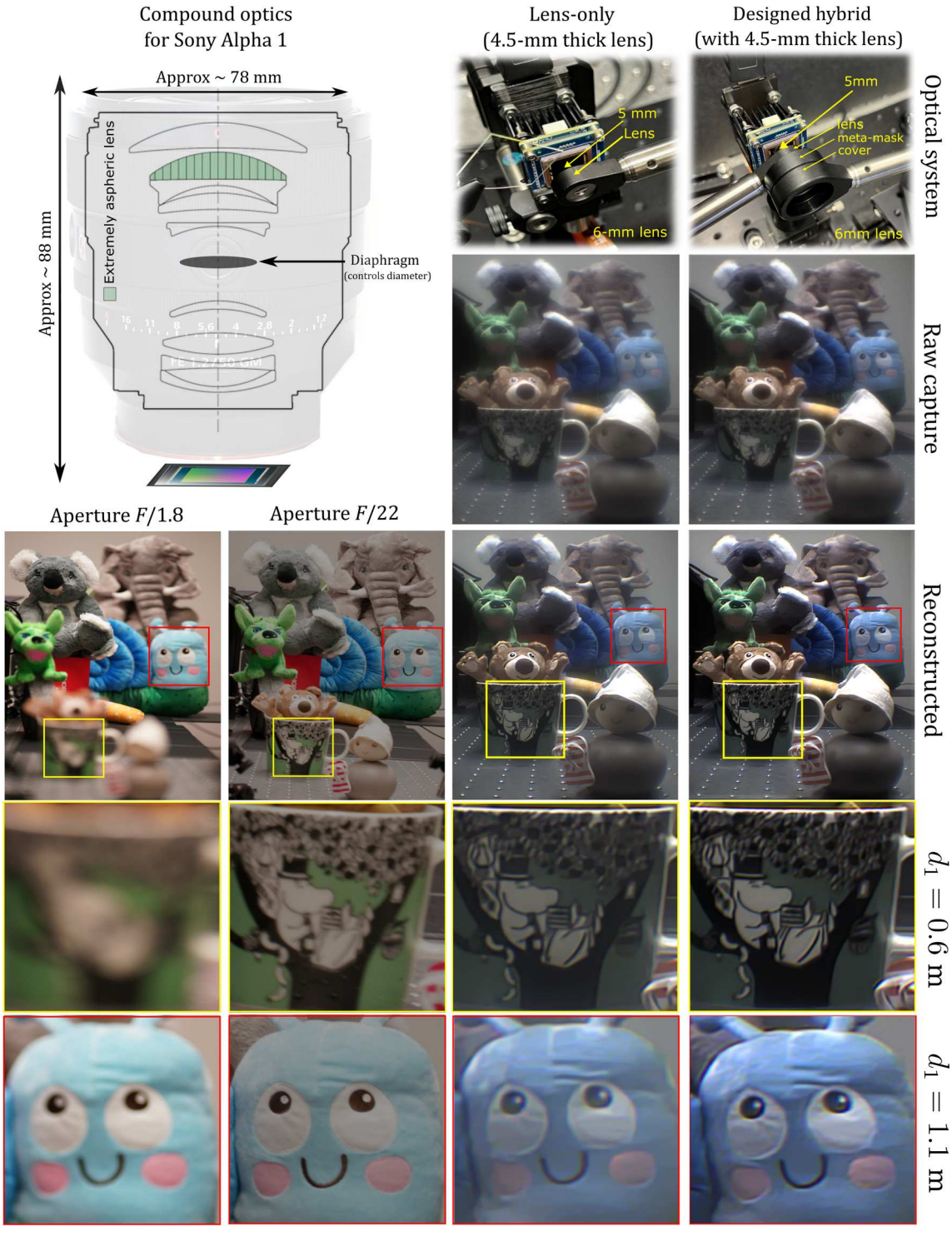

Fig. 4. Real-life scene composed of multiple objects of various colors arranged at different depths relative to the sensor. The top left shows a schematic of the compound lens, whereas top right images show the configuration for lens-only and hybrid systems. The second row (right columns) shows the raw capture of the hybrid and lens-only configuration. The reconstructed images are presented in row 3 with enlarged sections in rows 4 and 5, corresponding to off-focus distances of $d_1 = 0.6$ and 1.1 m, respectively. The results from the compound camera are obtained at the limit aperture values of *F*/1.8 and *F*/22.

compound lens, we demonstrate on-par or better performance. Key details about the developed hybrid camera are summarized in Table 1, which highlights that our designed hybrid meta-optics camera volume is approximately 108× thinner than a bulky commercial compound lens. An additional improvement of the designed hybrid meta-optics over the commercial compound lens and the lens-only setups is the extended depth of field (see Fig. 4). Essentially, by combining a refractive and meta-optics in one hybrid system, we circumvent otherwise stringent limitations that otherwise restrict the performance of either system alone. Specifically, no meta-optics has been demonstrated up to date, with an equivalent combination of image quality, large aperture size, low *F*-number, and depth-of-field (see the Supplementary Materials).

We envision that the developed design concept, where a HIL methodology is married with an end-to-end design framework, will be extended in future works toward diverse tasks, ranging from hyperspectral imaging, to classification or object detection tasks. This would leverage the full potential achievable with diffractive optics while mitigating the fundamental hurdle of the design process. Our optics design method takes an essential step toward ultra-small cameras that may enable novel applications in endoscopy, brain imaging, or in a distributed fashion on object surfaces.

MATERIALS AND METHODS

HIL optimizer

We used a nonlinear black-box optimizer that employs the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) strategy to optimize the phase profile at the SLM. This profile is modeled by Θ_{SLM} (see Fig. 1 and the Supplementary Materials). The loss function to be optimized by CMA-ES is the averaged PSNR between the target image and estimates from Convolutional Neural Networks (CNN)–based inverse imaging for three distances 0.5,1.0, and 1.8 m (see the setup in Fig. 3). Although CMA-ES is a 0th-order stochastic evolutionary method, it can be viewed as second-order since it estimates a covariance matrix closely related to the inverse Hessian (*29*). This feature allows CMA-ES to handle poorly conditioned problems. Methods that require derivatives like Adam (*30*) (frequently used to optimize optics in current state-of-the-art methodologies) directly contradict the HIL scenario. This contradiction is because the hardware is a "black box" of an unknown mathematical model and, as a result, is nondifferentiable. In addition, the efficiency of the CMA-ES method for black-box optimization of Image Signal Processor (ISP) is demonstrated by tuning a multi-objective highly nonlinear optimization problem in (*3*).

Table 1. Comparison of physical specifications between imaging systems. We remark that the total track length is the distance from the surface of the image sensor to the front surface of the first optics. Furthermore, the depth of field is the distance range in which the reconstructed image is sharp.

Features	Compound optics	Lens-only	Designed hybrid
Dimension	78.0 mm × 88.0 mm	5 mm × 9.5 mm	5 mm × 10.5 mm
Depth of field	0.8–2.0 m	0.8–1.2 m	0.3–2.0 m
Total track length	106 mm	9.5 mm	10.5 mm
Optics thickness	82 mm	4.5 mm	5.5 mm
F-number	*F*/22 – *F*/1.8	1.0	1.0

The HIL optimization is formalized using two models, one for the black-box camera and the second for the inverse imaging software. The observations registered by the sensor are defined as

$$y_i = \mathcal{T}_{\text{hard}}(\Theta_{\text{SLM}}, x_i) \tag{1}$$

where x_i is a true image, y_i represents the blurred image collected at the sensor, and Θ_{SLM} is the phase profile of the flat optics at the SLM. The operator $\mathcal{T}_{\text{hard}}$ in Eq. 1 includes image formation due to light-beam propagation through optical components of the system and signal transformations by other hardware elements in the loop. The mathematical model of this operator is unknown, and y_i corresponding to x_i can be found experimentally only. The varying phase profile of SLM Θ_{SLM} is subject to optimization.

On the other hand, the inverse imaging software operator $\mathcal{T}_{\text{soft}}$ reconstructs an estimate of x_i from y_i parametrized by Θ_{soft}. This process is modeled as

$$\hat{x}_i = \mathcal{T}_{\text{soft}}(\Theta_{\text{soft}}, y_i) \tag{2}$$

The design of hybrid camera requires to find the encoding phase modulation for DOE parametrized by Θ_{SLM} and the imaging software operator $\mathcal{T}_{\text{soft}}$ with parameter Θ_{soft}.

To design the flat optics and inverse imaging software jointly, we use an end-to-end approach using neural networks. To this end, the set of training images $\{x_i\}_{i=1}^{N}$ is given, the operator $\mathcal{T}_{\text{soft}}$ in Eq. 2 is defined by the structure of DRUNet (*33*), and Θ_{soft} are parameters of this network. This network can handle various noise levels for an RGB image, per channel, via a single model. The backbone of DRUNet is U-Net. Therefore, the goal of our end-to-end optimization is formulated as follows

$$(\hat{\Theta}_{\text{SLM}}, \hat{\Theta}_{\text{soft}}) \in arg\ min_{\Theta_{\text{SLM}}, \Theta_{\text{soft}}} \sum_{i=1}^{N} \mathcal{L}(\hat{x}_i, x_i) \tag{3}$$

where $\mathcal{L}$ is a loss function composed from peak PSNR values for accuracy evaluation and perceptual metrics (further details in the Supplementary Materials). Remark that combining Eqs. 1 and 2, we have that $\hat{x}_i = \mathcal{T}_{\text{soft}}[\Theta_{\text{soft}}, \mathcal{T}_{\text{hard}}(\Theta_{\text{SLM}}, x_i)]$ in Eq. 3. Once the hybrid camera is designed, the attained solution $(\hat{\Theta}_{\text{SLM}}, \hat{\Theta}_{\text{soft}})$ enables image reconstruction as

$$\hat{x}_i = \mathcal{T}_{\text{soft}}(\Theta_{\text{soft}} y_i) \tag{4}$$

Observe that the optics is fixed in Eq. 4, and it defines the inverse imaging for any y_i.

We wish to emphasize that this image reconstruction algorithm approximates the inverse imaging operator for image formation by $\mathcal{T}_{\text{hard}}$. The proposed inverse imaging software can be classified as a blind inverse imaging algorithm. This method is PSF-less which is cardinally different with respect to the conventional image reconstruction algorithms (*2*, *9*, *11*, *23*, *34*). This difference comes from the fact that they assume known or approximately known the model of $\mathcal{T}_{\text{hard}}$ and the corresponding PSFs.

Alternating iterations summarize the proposed design framework of hybrid optics to resolve the problem given by Eq. 3. First, fixing the phase hyperparameter of SLM Θ_{SLM}, solving Θ_{soft} for

inverse imaging. Second, updating Θ_{SLM} provide the fixed obtained Θ_{soft} and so forth with these alternating iterative optimizations. This optimization process starts by randomly selecting an SLM-pattern $\Theta_{SLM}^{(0)}$ or by an encoded phase profile derived at the stage of the model-based design. An inverse imaging CNN pretrained for a wide range of hyperparameters Θ_{SLM}. For optimization of Θ_{soft}, the algorithm acquires a set of blurred images obtained for an iterative value of $\Theta_{SLM}^{(r)}$ using the ISP of the optical system at distances from sensor 0.5,1.0, and 1.8 m. After, the acquired images are then passed to the downstream reconstruction module (deblurring using trained CNN). The output of the task module is evaluated by a domain-specific evaluation metric, which in this case is PSNR and perceptual metric as in Eq. 3.

For optimization of Θ_{SLM}, we use the 0th-order stochastic evolutionary search method CMA-ES (documentation in python of the CMA-ES optimizer in https://pypi.org/project/cma/) (*29*), and it updates Θ_{SLM}, taking advantage of the tested SLM patterns during the R iterations. Once Θ_{SLM} is updated, it refines the CNN-based inverse imaging by training it for the best $\Theta_{SLM}^{(r)}$. Performing the previous alternating process N_{iter} times algorithm returns the updated Θ_{SLM} and Θ_{soft}. The CNN structure used for inverse imaging (update on Θ_{soft}) is reported in Fig. 1. It can also be initiated by solutions obtained according to the model-based approach from (*35*).

The number of "global" iterations of this algorithm for Θ_{soft} is N_{iter}, and R is the number of updates (iterations) performed for Θ_{SLM}. The time needed to process physical observations is about 1.01 s for four hundred training images. This processing time involves the deconvolution of the blurred data collected at the sensor and the decision of the next SLM pattern. In this work, we fixed $N_{iter} = 3$ and $R = 500$.

Parameterized optics model

The phase profile of SLM is designed as a piece-wise spatially invariant function defined for the design wavelength λ_0. Following (*16*), we start from the absolute phase model. Further, this model is wrapped to the interval defined by the modulation range of SLM and discretized with a number of the parameters to be optimized equal to fourteen, which we introduce in brief. The proposed absolute phase ϕ_{λ_0} takes the form

$$\varphi_{\lambda_0}(x,y) = \frac{2\pi\alpha}{\lambda_0 f_{\lambda_0}}(x^2+y^2) + \beta(x^3+y^3) + \sum_{p=1,p=4}^{P} \rho_p Z_p(x,y) \quad (5)$$

where the first term models the squared phase of the lens which fits the form of the fourth Zernike polynomial omitted in the sum of the Zernike components, α encapsulates the focusing contribution of the SLM for wavelength λ_0, and the lens focal length f_{λ_0} fixed as $f_{\lambda_0} = 5$ mm in this work. The cubic phase of magnitude β is a typical component to extend the depth of field as introduced in (*36*). The third group of the items is for parametric approximation of the free-shape DOE using the Zernike polynomials $Z_p(x, y)$ with coefficients ρ_p to be estimated. Thus, the complete set of parameters to be optimized for SLM is defined as $\Theta_{SLM} = (\alpha, \beta, \rho_1, ..., \rho_P)$.

Fresnel order (thickness)

The SLM phase profile in radians is defined as $Q = 2\pi m_Q$, where m_Q is called the "Fresnel order" of the mask, which in general is not necessarily an integer. Then, the phase profile of SLM considering the thickness is calculated as $\hat{\varphi}_{\lambda_0}(x,y) = \mathrm{mod}\left[\varphi_{\lambda_0}(x,y) + \frac{Q}{2}, Q\right] - Q/2$.This operation returns $\hat{\varphi}_{\lambda_0}(x,y)$ taking the values in the interval $[-Q/2, Q/2)$. The parameter m_Q is known as the Fresnel order of phase pattern. For $m_Q = 1$, this restriction to the interval $[-\pi, \pi)$ corresponds to the standard phase wrapping operation. For the experiments, we will present that, in the next section, m_Q is fixed to 1, providing the maximum phase thickness that our SLM supports, e.g., phase interval $[-\pi, \pi)$.

Piecewise invariant phase

The SLM pattern is defined on 2D grid (X, Y) with the computational sampling period (computational pixel) Δ_{comp}. We obtain a piece-wise invariant surface for SLM after a nonlinear transformation of the absolute phase. The discrete uniform grid of the wrap phase profile $\hat{\varphi}_{\lambda_0}(x,y)$ to the N levels is given as $\theta_{\lambda_0}(x,y) = \lfloor \hat{\varphi}_{\lambda_0}(x,y)/N \rfloor \cdot N$, where $\lfloor w \rfloor$ stays for the integer part of w. The values of $\theta_{\lambda_0}(x, y)$ are restricted to the interval $[-Q/2, Q/2)$. Q is an upper bound for thickness phase of $\theta_{\lambda_0}(x, y)$.

Let m_w be the size of SLM's pixel, and let a computational pixel Δ_{comp} be larger than m_w. The latter assumption is natural as SLM cannot implement smaller values of Δ_{comp}. In this design, Δ_{comp} serves as a lower bound for steps of the SLM phase pattern. Larger Δ_{comp} means a more straightforward structure of phase pattern and a simpler corresponding phase mask for implementation. Fresnel order Q, the number of steps N, and computational pixel Δ_{comp} could be included in Θ_{SLM} as design parameters.

Meta-surface design and implementation

To build the compact miniaturized color camera, we created a meta-optics in SiN (*31*), taking advantage of its low cost and straightforward manufacturability. Specifically, the described optimized phase mask (see Fig. 2) was implemented through square pillars with a height of 700 nm, a period of 350 nm, and varying side widths at a design wavelength of 540 nm. For this purpose, we first calculated the phase and amplitude response using rigorous coupled wave analysis (*32*) (further detailed in the Supplementary Materials). From the resulting width phase-delay relation, we chose discrete levels of the scatterer mapped to the desired phase profile. We fabricated the meta-optics using a nanofabrication approach in a 700-nm PECVD-grown SiN on a quartz substrate (SPTS Delta LPX). The geometry was first patterned in a polymer resist (ZEP-520A) via electron beam lithography (JEOL JBX-6300FS, 100*kV*). Then, a hard alumina mask was deposited using thermal evaporation. After lift-off, the pattern was transferred into the underlying SiN layer using reactive ion etching (Oxford Instruments, PlasmaLab 100) with fluorine etch chemistry (SF6/C4F8).

Supplementary Materials

This PDF file includes:

Supplementary Text
Figs. S1 to S12
References

REFERENCES AND NOTES

1. L. Huang, S. Colburn, A. Zhan, A. Majumdar, Full-color metaoptical imaging in visible light. *Adv. Photon. Res.* **3**, 2100265 (2022).
2. E. Tseng, S. Colburn, J. Whitehead, L. Huang, S. Baek, A. Majumdar, F. Heide, Neural nano-optics for high-quality thin lens imaging. *Nat. Commun.* **12**, 6493 (2021).

3. S. Pinilla, S. Rostami, I. Shevkunov, V. Katkovnik, K. Egiazarian, Hybrid diffractive optics design via hardware-in-the-loop methodology for achromatic extended-depth-of-field imaging. *Opt. Express* **30**, 32633–32649 (2022).
4. W. Chen, A. Zhu, J. Sisler, Y. Huang, K. Yousef, E. Lee, C. Qiu, F. Capasso, Broadband achromatic metasurface-refractive optics. *Nano Lett.* **18**, 7801–7808 (2018).
5. A. Zhan, S. Colburn, C. Dodson, A. Majumdar, Metasurface freeform nanophotonics. *Sci. Rep.* **7**, 1673 (2017).
6. R. Sawant, D. Andrén, R. Martins, S. Khadir, R. Verre, M. Käll, P. Genevet, Aberration-corrected large-scale hybrid metalenses. *Optica* **8**, 1405–1411 (2021).
7. V. Sitzmann, S. Diamond, Y. Peng, X. Dun, S. Boyd, W. Heidrich, F. Heide, G. Wetzstein, End-to-end optimization of optics and image processing for achromatic extended depth of field and super-resolution imaging. *ACM Trans. Graph.* **37**, 114 (2018).
8. J. Chang, V. Sitzmann, X. Dun, W. Heidrich, G. Wetzstein, Hybrid optical-electronic convolutional neural networks with optimized diffractive optics for image classification. *Sci. Rep.* **8**, 12324 (2018).
9. X. Dun, H. Ikoma, G. Wetzstein, Z. Wang, X. Cheng, Y. Peng, Learned rotationally symmetric diffractive achromat for full-spectrum computational imaging. *Optica* **7**, 913–922 (2020).
10. S. Baek, H. Ikoma, D. Jeon, Y. Li, W. Heidrich, G. Wetzstein, M. Kim, End-to-end hyperspectral-depth imaging with learned diffractive optics. ArXiv:2009.00463 (2020).
11. H. Arguello, S. Pinilla, Y. Peng, H. Ikoma, J. Bacca, G. Wetzstein, Shift-variant color-coded diffractive spectral imaging system. *Optica* **8**, 1424–1434 (2021).
12. S. Colburn, A. Majumdar, Inverse design and flexible parameterization of meta-optics using algorithmic differentiation. *Commun. Phys.* **4**, 65 (2021).
13. Z. Lin, C. Roques-Carmes, R. Pestourie, M. Solja˘ić, A. Majumdar, S. Johnson, End-to-end nanophotonic inverse design for imaging and polarimetry. *Nanophotonics* **10**, 1177–1187 (2021).
14. A. Mosleh, A. Sharma, E. Onzon, F. Mannan, N. Robidoux, F. Heide, Hardware-in-the-Loop End-to-End Optimization of Camera Image Processing Pipelines,in *Proceedings Of The IEEE/CVF Conference On Computer Vision And Pattern Recognition (CVPR)* (2020), pp. 7529–7538.
15. J. E. M. Whitehead, A. Zhan, S. Colburn, L. Huang, A. Majumdar, Fast extended depth of focus meta-optics for varifocal functionality. *Photonics Res.* **10**, 828–833 (2022).
16. S. R. M. Rostami, V. Katkovnik, K. Eguiazarian, Extended DoF and achromatic inverse imaging for lens and lensless MPM camera based on Wiener filtering of defocused OTFs. *Opt. Eng.* **60**, 14 (2021).
17. Q. Fan, W. Xu, X. Hu, W. Zhu, T. Yue, C. Zhang, F. Yan, L. Chen, H. Lezec, Y. Lu, A. Agrawal, T. Xu, Others Trilobite-inspired neural nanophotonic light-field camera with extreme depth-of-field. *Nat. Commun.* **13**, 2130 (2022).
18. O. Lévêque, C. Kulcsár, A. Lee, H. Sauer, A. Aleksanyan, P. Bon, L. Cognet, F. Goudail, Co-designed annular binary phase masks for depth-of-field extension in single-molecule localization microscopy. *Opt. Express* **28**, 32426–32446 (2020).
19. F. Presutti, F. Monticone, Focusing on bandwidth: Achromatic metalens limits. *Optica* **7**, 624–631 (2020).
20. Q. Sun, C. Wang, F. Qiang, D. Xiong, H. Wolfgang, End-to-end complex lens design with differentiable ray tracing. *ACM Trans. Graph.* **40**, 1–13 (2021).
21. Y. Liu, C. Zhang, T. Kou, Y. Li, Y. J. Shen, End-to-end computational optics with a singlet lens for large depth-of-field imaging. *Opt. Express* **29**, 28530–28548 (2021).
22. J. Tan, L. Niu, J. Adams, V. Boominathan, J. Robinson, R. Baraniuk, A. Veeraraghavan, Face detection and verification using lensless cameras. *IEEE Trans. Comput. Imaging* **5**, 180–194 (2019).
23. D. Jeon, S. Baek, S. Yi, Q. Fu, X. Dun, W. Heidrich, M. Kim, Compact snapshot hyperspectral imaging with diffracted rotation. *ACM Transact. Graphics* **38**, 117 (2019).
24. M. Asif, A. Ayremlou, A. Sankaranarayanan, A. Veeraraghavan, R. Baraniuk, FlatCam: Thin, lensless cameras using coded aperture and computation. *IEEE Trans. Comput. Imaging* **3**, 384–397 (2017).
25. K. Monakhova, K. Yanny, N. Aggarwal, L. Waller, Spectral DiffuserCam: Lensless snapshot hyperspectral imaging with a spectral filter array. *Optica* **7**, 1298–1307 (2020).
26. S. Colburn, A. Zhan, A. Majumdar, Metasurface optics for full-color computational imaging. *Sci. Adv.* **4**, eaar2114 (2018).
27. S. MiriRostami, S. Pinilla, I. Shevkunov, V. Katkovnik, K. Egiazarian, Hybrid diffractive optics (DOE & refractive lens) for broadband EDoF imaging. *Electron. Imag.* **35**, 1–14 (2023).
28. V. Katkovnik, I. Shevkunov, N. Petrov, K. Egiazarian, Computational super-resolution phase retrieval from multiple phase-coded diffraction patterns: Simulation study and experiments. *Optica* **4**, 786–794 (2017).
29. N. Hansen, A. Ostermeier, Adapting Arbitrary Normal Mutation Distributions in Evolution Strategies: The Covariance Matrix Adaptation, in *Proceedings Of IEEE International Conference On Evolutionary Computation* (1996), pp. 312–317.
30. D. Kingma, J. Ba, Adam: A method for stochastic optimization. ArXiv Preprint ArXiv:1412.6980 (2014).
31. A. Zhan, S. Colburn, R. Trivedi, T. Fryett, C. Dodson, A. Majumdar, Low-contrast dielectric metasurface optics. *ACS Photonics* **3**, 209–214 (2016).
32. V. Liu, S. Fan, S^4: A free electromagnetic solver for layered periodic structures. *Comput. Phys. Commun.* **183**, 2233–2244 (2012).
33. K. Zhang, Y. Li, W. Zuo, L. Zhang, L. Van Gool, Timofte, R. Plug-and-play image restoration with deep denoiser prior, in *IEEE Transactions On Pattern Analysis And Machine Intelligence* (2022).
34. H. Hu, H. Zhou, Z. Xu, Q. Li, H. Feng, Y. Chen, T. Jiang, W. Xu, Practical snapshot hyperspectral imaging with DOE. *Opt. Lasers Eng.* **156**, 107098 (2022).
35. S. Rostami, S. Pinilla, I. Shevkunov, V. Katkovnik, K. Egiazarian, Power-balanced hybrid optics boosted design for achromatic extended depth-of-field imaging via optimized mixed OTF. *Appl. Optics* **60**, 9365–9378 (2021).
36. E. Dowski, W. Cathey, Extended depth of field through wave-front coding. *Appl. Optics* **34**, 1859–1866 (1995).
37. K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition. ArXiv Preprint ArXiv:1409.1556 (2014).
38. S. S. Khan, V. Sundar, V. Boominathan, A. Veeraraghavan, K. Mitra, Flatnet: Towards photorealistic scene reconstruction from lensless measurements. *IEEE Trans. Pattern Anal. Mech. Intell.* **44**, 1934–1948 (2020).
39. I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* **27**, (2014).
40. P. Ramachandran, B. Zoph, Q. Le, Searching for activation functions. ArXiv Preprint ArXiv:1710.05941 (2017).
41. S. Rostami, S. Pinilla, I. Shevkunov, V. Katkovnik, K. Egiazarian, On design of hybrid diffractive optics for achromatic extended depth-of-field (EDoF) RGB imaging. ArXiv Preprint ArXiv:2203.16985 (2022).

Acknowledgments

Funding: This work is supported by the CIWIL project funded by Jane and Aatos Erkko Foundation, Finland. I.S. was supported by the Academy of Finland (project no. 336357, PROFI 6–TAU Imaging Research Platform). Part of this work was conducted at the Washington Nanofabrication Facility/Molecular Analysis Facility, a National Nanotechnology Coordinated Infrastructure (NNCI) site at the University of Washington, with partial support from the National Science Foundation via awards NNCI-1542101, NNCI-2025489, and DARPA W31P4Q21C0043.

Author contributions: Conceptualization: S.P., V.K., and I.S. Methodology: S.P., J.E.F., and S.R.M.R. Investigation: S.P., S.R.M.R., and J.E.F. Visualization: S.P. Supervision: K.E. and A.M. Writing—original draft: All coauthors. Writing—review and editing: All coauthors. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials.

Submitted 17 January 2023
Accepted 20 April 2023
Published 26 May 2023
10.1126/sciadv.adg7297

ScienceAdvances

Miniature color camera via flat hybrid meta-optics

Samuel Pinilla, Johannes E. Frch, Seyyed Reza Miri Rostami, Vladimir Katkovnik, Igor Shevkunov, Arka Majumdar, and Karen Egiazarian

Sci. Adv., **9** (21), eadg7297.
DOI: 10.1126/sciadv.adg7297