



OPTICS

Real-time machine learning-enhanced hyperspectro-polarimetric imaging via an encoding metasurface

Lidan Zhang†, Chen Zhou†, Bofeng Liu†, Yimin Ding, Hyun-Ju Ahn, Shengyuan Chang, Yao Duan, Md Tarek Rahman, Tunan Xia, Xi Chen, Zhiwen Liu*, Xingjie Ni*

Light fields carry a wealth of information, including intensity, spectrum, and polarization. However, standard cameras capture only the intensity, disregarding other valuable information. While hyperspectral and polarimetric imaging systems capture spectral and polarization information, respectively, in addition to intensity, they are often bulky, slow, and costly. Here, we have developed an encoding metasurface paired with a neural network enabling a normal camera to acquire hyperspectro-polarimetric images from a single snapshot. Our experimental results demonstrate that this metasurface-enhanced camera can accurately resolve full-Stokes polarization across a broad spectral range (700 to 1150 nanometer) from a single snapshot, achieving a spectral sensitivity as high as 0.23 nanometer. In addition, our system captures full-Stokes hyperspectro-polarimetric video in real time at a rate of 28 frames per second, primarily limited by the camera's readout rate. Our encoding metasurface offers a compact, fast, and cost-effective solution for multi-dimensional imaging that effectively uses information within light fields.

INTRODUCTION

The ability to observe and use multidimensional information of light has been a long-standing human pursuit, offering a more comprehensive understanding of light-matter interaction and promoting more accurate characterizations across many fields (1–4). For instance, the spectral information can significantly improve the precision of medical diagnosis and agricultural monitoring (5, 6), while the polarization information aids in material classification and stress analysis (7). Obtaining multidimensional information within a single system presents substantial advantages.

However, current hyperspectral and polarimetric imaging techniques often capture these two distinct types of data separately, making it difficult to rapidly obtain comprehensive information (8–11). Furthermore, conventional spectral and polarization imaging systems, which predominantly rely on diffractive optics or optical filters, tend to be bulky and heavy. (2) These optical elements are typically sensitive to either the spectrum or the polarization of incoming light. A common solution involves the incorporation of polarization elements into spectral systems. For example, stacking layers of organic filters can create integrated hyperspectro-polarimetric (HSP) cameras at the micrometer scale (12). Another method entails integrating linear micropolarizer arrays onto traditional red, green, and blue Bayer sensor (13, 14). However, both of those approaches have limited number of wavelength channels and are only capable of analyzing linear polarization states. These limitations not only confine the imaging systems' capacity to incorporate multiple functionalities into a single platform but also obstruct their potential for further miniaturization and integration (8, 15–17).

Recently, a few HSP cameras have been developed leveraging optical metasurfaces. A metasurface is a thin layer of nanostructures capable of tailoring various light properties, including polarization,

amplitude, and phase, presenting a paradigm shift in miniaturizing optical devices. It enables various applications such as beam bending, beam shaping, holography, and even invisibility cloaks (18–22). One type of spectro-polarimetric camera relies on a spatial-multiplexing metasurface lens that directs light of different polarizations and wavelengths onto separate focal points (23). However, the method suffers from cross-talk among different polarization and spectral channels, and it has limited spectral resolution due to the inherent coupling of the spectral and spatial dimensions. Another approach uses multiple rotating metasurfaces but is limited to linear polarization analysis. Moreover, it involves mechanically moving parts, constraining the system's compactness and making real-time imaging considerably challenging (24).

Here, we present a metasurface-enhanced camera capable of capturing HSP images in a single shot and videos in real time. This camera obtains comprehensive information about the broadband spectrum and all four Stokes parameters simultaneously. The core technology lies in our uniquely designed metasurface encoder, which encodes spectral and polarization information into a spatial intensity distribution that can be captured by a standard camera. This is paired with a custom machine learning (ML) neural network (NN) decoder to enable real-time image recovery. Note that the concept of using broadband filters has been introduced to realize a novel spectrometer (10, 25–29). In contrast to the use of narrowband filters or gratings in a conventional spectrometer, the wavelength information is mingled by those filters and then resolved in the postprocessing, given that the exact spectral responses of the filters are known (30). Our metasurface is designed such that each superpixel has both distinct spectral and polarization responses. Specifically, we have incorporated chiral meta-atoms exhibiting strong chirality, enabling different polarization responses for all polarization states, including both left- and right-circular polarizations. Coupling the metasurface encoder with our ML NN decoder, we achieved real-time acquisition of HSP images at a rate of 28 frames per second (FPS), mainly limited by our camera's maximal readout speed. Our device operates within the wavelength

Department of Electrical Engineering, the Pennsylvania State University, University Park, PA 16802, United States.

*Corresponding author. Email: zzl1@psu.edu (Z.L.); xingjie@psu.edu (X.N.)

†These authors contributed equally to this work.

range from 700 to 1150 nm, among the highest bandwidths reported for integrated spectrum polarization devices (12, 23, 31, 32). It reveals a high spectral sensitivity of down to 0.23 nm and extracts complete Stokes polarization information in addition to the wavelength data. We have demonstrated the device's ability to distinguish our test images with arbitrary wavelength and polarization distribution, as well as natural objects such as beetles, given its capability to acquire spectral and complete polarization data simultaneously. We envision that our real-time HSP camera can greatly enhance imaging systems functioning in dynamic, high-speed environments such as surveillance systems, autonomous vehicles, or biomedical systems. This advancement can be leveraged in diverse fields, including biodiversity research, behavioral studies, and environmental monitoring.

RESULTS

Our metasurface features thousands of spectro-polarimetric superpixels, each comprising arrays of judiciously designed chiral meta-atoms. The metasurface was placed in conjunction of the detector chip of a standard monochromatic camera, enabling it to image objects with both wavelength and polarization data. In contrast to traditional narrow-band filter-based spectrometers or single-channel polarizers—where each wavelength or polarization state is assigned to a single pixel in the spatial domain—our metasurface multiplexes all spectro-polarimetric channels through a single superpixel (note S1). Each array of meta-atoms acts as a unique polarization and wavelength encoder, with $h(\lambda, S)$ denoting its transmission function. By combining multiple arrays with distinct responses, we can establish the spectro-polarimetry that functions as an image pixel for our HSP camera. The intensity information $I(x, y)$ captured by the camera is subsequently decoded by a ML backend, producing four-dimensional (4D) images $f'(x, y, \lambda, S)$ of the input. Figure 1 shows an artistic depiction of such a 4D image, with distinct wavelength and polarization views, each emphasizing different features of the objects.

The encoding process of spectro-polarimetry by the metasurface can be elucidated using a mathematical model. The wavelength and polarization information are encoded such that the intensity received by each camera pixel is the integral of the entire spectrum, weighted by the spectral and polarization response of the corresponding meta-atoms directly above it. We describe the polarization state with the Stokes parameters $S = [S_0, S_1, S_2, S_3]$. Thus, the transmitted intensity received by camera pixels behind the i -th spectro-polarization encoder can be expressed as

$$I_i = \sum_{k=0}^3 \int_{\lambda_{\min}}^{\lambda_{\max}} T_{iS_k}(\lambda) f_{S_k}(\lambda) d\lambda \quad (1)$$

where $T_{iS_0}(\lambda)$, $T_{iS_1}(\lambda)$, $T_{iS_2}(\lambda)$, and $T_{iS_3}(\lambda)$ are the transmission coefficients for different S components at the incident wavelength λ , and $f_{S_0}(\lambda)$, $f_{S_1}(\lambda)$, $f_{S_2}(\lambda)$, and $f_{S_3}(\lambda)$ are the different S components for the light signal at the wavelength λ . Muller matrix is widely used for manipulating S components and here $T_{iS_0}(\lambda)$, $T_{iS_1}(\lambda)$, $T_{iS_2}(\lambda)$, and $T_{iS_3}(\lambda)$ are the first row of the Muller matrix at wavelength λ . The transmission equation can be written in a matrix form as

$$\begin{aligned} I &= T f_s \\ I &= (I_1, \dots, I_n)^T \\ f_s &= (f_{1S_0}, f_{1S_1}, f_{1S_2}, f_{1S_3}, \dots, f_{mS_0}, f_{mS_1}, f_{mS_2}, f_{mS_3})^T \\ T &= \begin{pmatrix} T_{11}^{\lambda_{1S_0}} & T_{11}^{\lambda_{1S_1}} & T_{11}^{\lambda_{1S_2}} & T_{11}^{\lambda_{1S_3}} & \dots & T_{1m}^{\lambda_{1S_0}} & T_{1m}^{\lambda_{1S_1}} & T_{1m}^{\lambda_{1S_2}} & T_{1m}^{\lambda_{1S_3}} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ T_{n1}^{\lambda_{mS_0}} & T_{n1}^{\lambda_{mS_1}} & T_{n1}^{\lambda_{mS_2}} & T_{n1}^{\lambda_{mS_3}} & \dots & T_{nm}^{\lambda_{mS_0}} & T_{nm}^{\lambda_{mS_1}} & T_{nm}^{\lambda_{mS_2}} & T_{nm}^{\lambda_{mS_3}} \end{pmatrix} \end{aligned} \quad (2)$$

where n represents the number of detectors and m denotes the number of wavelength sampling points. Although mathematically the

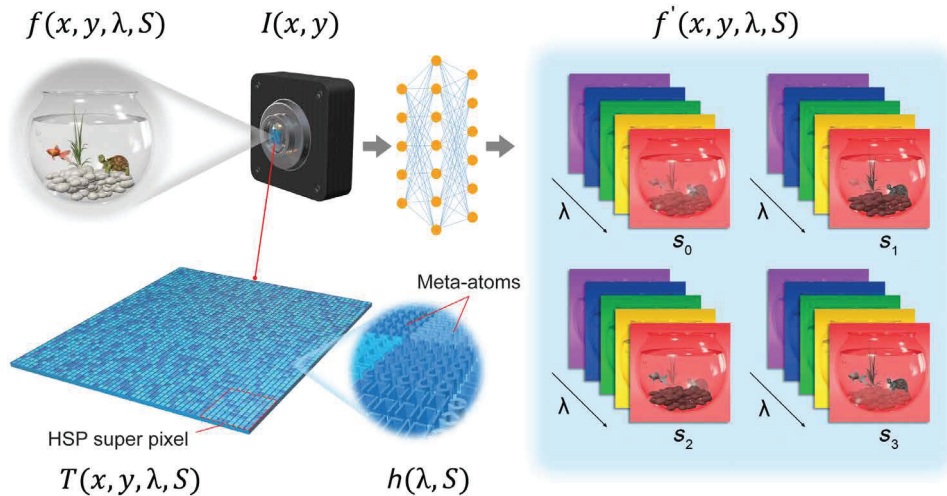


Fig. 1. Conceptual diagram of the proposed HSP camera. The image of an object (e.g., a fish tank), which has spectral and polarization information denoted by $f(x, y, \lambda, S)$, is transformed into a 2D intensity mapping, $I(x, y)$, by the HSP camera. This HSP camera system comprises two main components: encoding metasurfaces at the frontend to encode the information and a ML algorithm at the backend to decode the information. Through the HSP camera, the encoded 2D intensity information of the tank image is expanded into 4D information $f'(x, y, \lambda, S)$. Certain parts of this image (such as the fish and the rock) in S_2 become more pronounced when viewed under different polarizations. The encoding metasurfaces encompass multiple wavelength and polarization encoders $h(\lambda, S)$, which can be flexibly combined to construct an HSP superpixel $T(x, y, \lambda, S)$ according to the required resolution.

spectral and polarization information can be recovered by directly pseudo-inverting the transmission matrix T , the recovery may be compromised, leading to inaccurate results due to various unavoidable errors and uncertainties, such as calibration noise, measurement noise, and nonuniform responses of the camera pixels, etc. To address this, we first used linear regression with the least-square optimization method (33) to retrieve spectral and polarization information, which can be formulated as

$$\begin{aligned} \hat{f}_s &= \underset{f_s}{\operatorname{argmin}} \|I_{\text{meas}} - T f_s\|^2 \\ \text{Subject to } f_{is_0} &\geq 0, f_{is_0}^2 \geq f_{is_1}^2 + f_{is_2}^2 + f_{is_3}^2 \end{aligned} \quad (3)$$

The objective is to minimize the discrepancy between the measured intensity I_{meas} and the intensity $T f_s$ calculated from the transmission matrix. It subjects to both linear and nonlinear constraints because the Stokes parameters must satisfy the inequities as shown in Eq. 3. These constraints introduce nonconvexity to the problem, making it challenging to solve (34). It can be addressed using local optimization methods such as the interior-point method (35, 36) or using global optimization strategies such as particle swarm algorithms (37). While the latter might produce better images, it comes at the expense of increased computational time. To further enhance image quality, advanced algorithms incorporating sparsity (38) in both time and space dimensions can be used. However, this approach is considerably more time intensive. Therefore, we developed a method based on the ML NN to perform fast decoding of spectral and polarization information.

ML-based fast decoding

ML can approximate complex functions given sufficient training data, and it does not necessitate an explicit physical model. Here, we used a customized ML network to reconstruct HSP images accurately and quickly. Given training data, the model can learn the mapping relationship between an encoded input image and its corresponding wavelength and polarization information. In our case, each superpixel in the captured image corresponds to a unique spectrum and polarization information, independent of the other pixels.

For traditional optimization methods, a unique model is required for each superpixel due to fabrication-induced nonuniformity across the superpixels and optical aberrations of the system. As a result, thousands of models are required, leading to high computational costs. In contrast, with the proposed ML approach, we establish a unified network that takes in both the position encodings of the superpixels and the encoded intensity distributions for polarization and spectrum reconstruction. This way, all superpixels are processed using a single network. Because the spatial locations of the superpixels are also part of the input, the network can learn the effects of the nonuniformity and aberrations and compensate for it. Once the ML model is trained, the decoding process only requires simple calculations, making possible real-time recovery of HSP images.

Design and characterization of the metasurface

The metasurface for our HSP camera is composed of thousands of spectro-polarimetric encoders (superpixel), each consisting of an array of meta-atoms (Fig. 2A). For discerning the multichannel information, each meta-atom array exhibits a unique response to different wavelengths and polarizations. The meta-atoms are silicon based and

have shapes of split rings or split doors. These designs manifest not only anisotropy but also strong chirality. For example, the split ring design features two unequal splits and arm lengths. The coupling between the bright mode (the electric dipole resonance supported by the arms) and the dark mode (the magnetic dipole resonance supported by the entire ring) leads to a sharp Fano resonance, essential for achieving high spectral resolution. In addition, the split ring also exhibits strong chirality (39), vital for distinguishing left- and right-handed elliptical/circular polarizations. We established a large library of meta-atoms by varying the structural parameters of the split ring and split door designs while keeping their height and period constant (see note S2). We performed numerical simulations using a commercial finite element method solver package, COMSOL Multiphysics, to calculate the spectral response with the x and y linear polarization inputs, respectively. Each meta-atom is modeled in a periodic array where the period is chosen to give at least 80-nm distance between neighboring meta-atoms to avoid strong near-field coupling. The transmission of the meta-atoms for different polarization and different wavelength were used to obtain the Jones matrix, which was then converted to the Muller matrix. Figure 2B shows a typical response from a meta-atom, showcasing rich spectral features for different polarization inputs. This highlights its capability to distinguish full-Stokes polarization and wavelength information. For added diversity to the library, we incorporated structures with identical geometric parameters but modified by either a 90° rotation, a mirrored orientation, or both. In total, our meta-atom library comprises 1936 split-door shapes and 1294 split-ring shapes.

It is essential to select meta-atom designs that exhibit distinct responses to different wavelengths and polarizations, as the dissimilarity between the responses of the structures determines the resolving power of the reconstructive spectrometer polarimetry. If two designs produce similar responses, then it becomes challenging to discern the wavelength and polarization. Mathematically, the response of each meta-atom design is represented by a column in the transmission matrix, T , described in Eq. 2. Consequently, the condition number of the transmission matrix indicates the similarity of the responses. We chose 100 designs from the meta-atom library to construct the metasurface so that the resulting transmission matrix has a minimum condition number (as shown in Fig. 2C). We arranged each meta-atom design in a $9\ \mu\text{m} \times 9\ \mu\text{m}$ 2D array (Fig. 2A). In total, there are 100 different arrays. We then organized these arrays into a 2D grid to construct a $90\ \mu\text{m} \times 90\ \mu\text{m}$ superpixel. Each metasurface consists of 20×20 superpixels, which are also arranged in a 2D grid. When aiming for different spatial, spectral, and polarization resolution, the superpixel size and number can be adjusted accordingly. After fabricating our metasurface (see the detailed fabrication process in Materials and Methods), we calibrated the transmission of the pixels under a variety of polarization and wavelength combinations (see the detailed setup in fig. S3). Figure 2D shows the results of 30 different wavelengths and 6 incident polarizations. We observed that different wavelengths and polarizations exhibit distinct intensity distribution patterns, as captured by the camera (DMK 33GX265). These patterns were then used to determine the transmission matrix or for training the ML NN.

Optimization-based recovery

We first demonstrated the capability of our HSP camera using optimization-based recovery. By knowing the captured intensity distribution along with its corresponding input, we calculated the

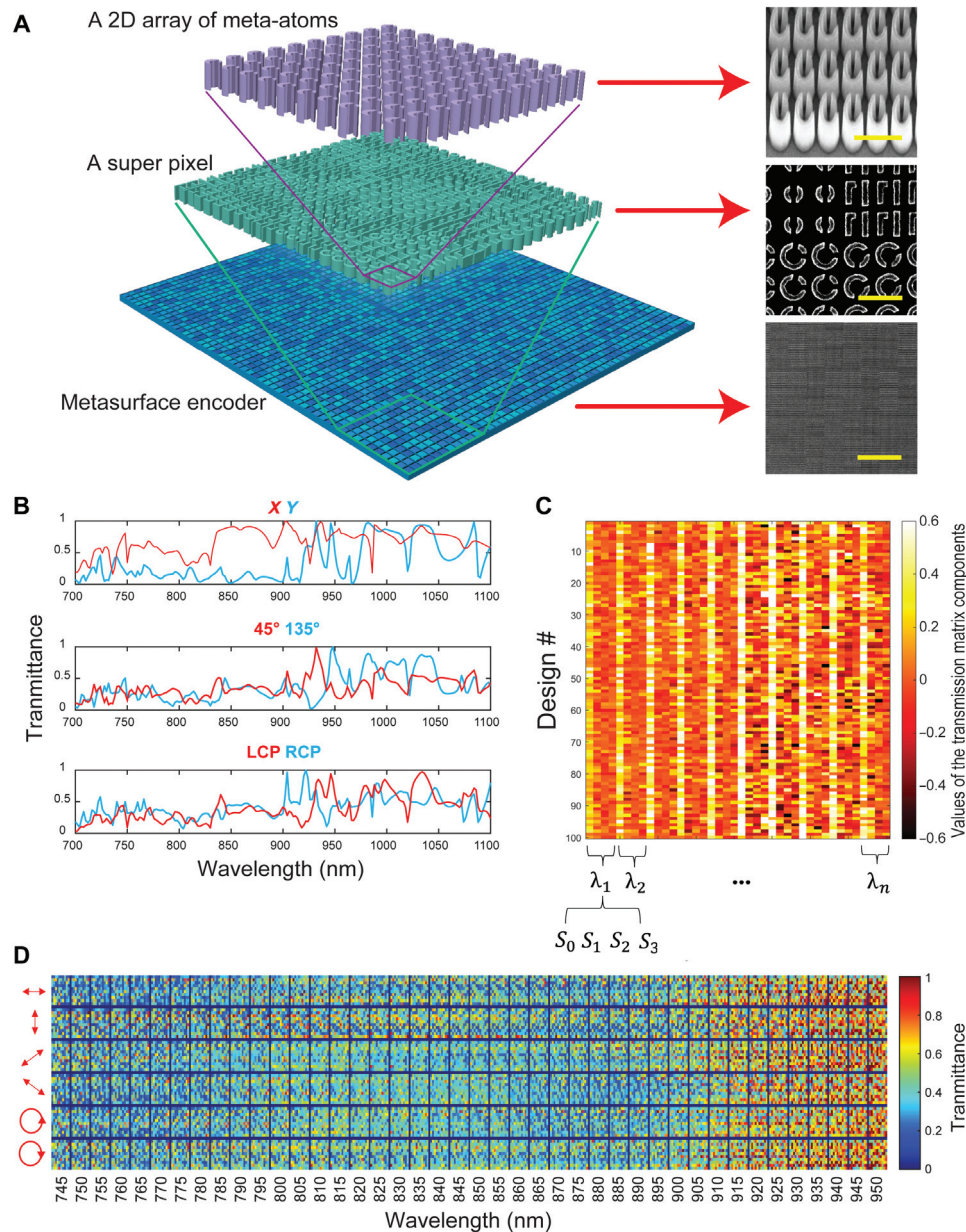


Fig. 2. Design and characterization of the encoding metasurfaces. (A) Left panel shows an artistic illustration of the metasurface, and the right panels show the SEM images of the fabricated metasurface. Scale bars, 1 μm , 1 μm , and 200 μm (from top right to bottom right). (B) Spectral response for different polarization inputs for a typical split-ring-shaped meta-atom. (C) The transmission matrix for our selected 100 chiral meta-atoms, with the condition number (CN) minimized to 37.15, indicating good distinguishability of different meta-atom's spectral and polarization responses. (D) Each small box represents the intensity response of a single superpixel, which consists of a 10×10 array of meta-atoms, as depicted in the central part of the right panel of (A). The intensity responses under different wavelength and polarization illuminations captured using a regular camera. For demonstration purposes, the intensity response of one typical superpixel is shown here.

transmission matrix, T , of each superpixel based on Eq. 2. Using the transmission matrix, we demonstrated that the spectral polarimetry of unknown incident beams could be retrieved through the optimization-based recovery method described in Eq. 3.

We prepared an input beam covering a broad wavelength range by combining two laser beams: one with left circular polarization (LCP) at 800 nm and the other with x polarization at 1100 nm. After passing through our metasurface, the encoded patterns were captured by the camera. Figure 3A shows the recovered Stokes

spectra ranging from 700 to 1150 nm, aligning well with those measured directly by a spectrometer with Stokes analysis (fig. S4). We demonstrated that our device operates over a bandwidth of 450 nm, which is limited by the range of our tunable laser source used in the experiment.

Furthermore, to achieve high spectral resolution, we calibrated the metasurface with a spectral step size of 0.23 nm, limited by the spectrometer used. We introduced an input beam with wavelengths ranging from 760 to 776 nm, with x polarization at 761.8 nm and

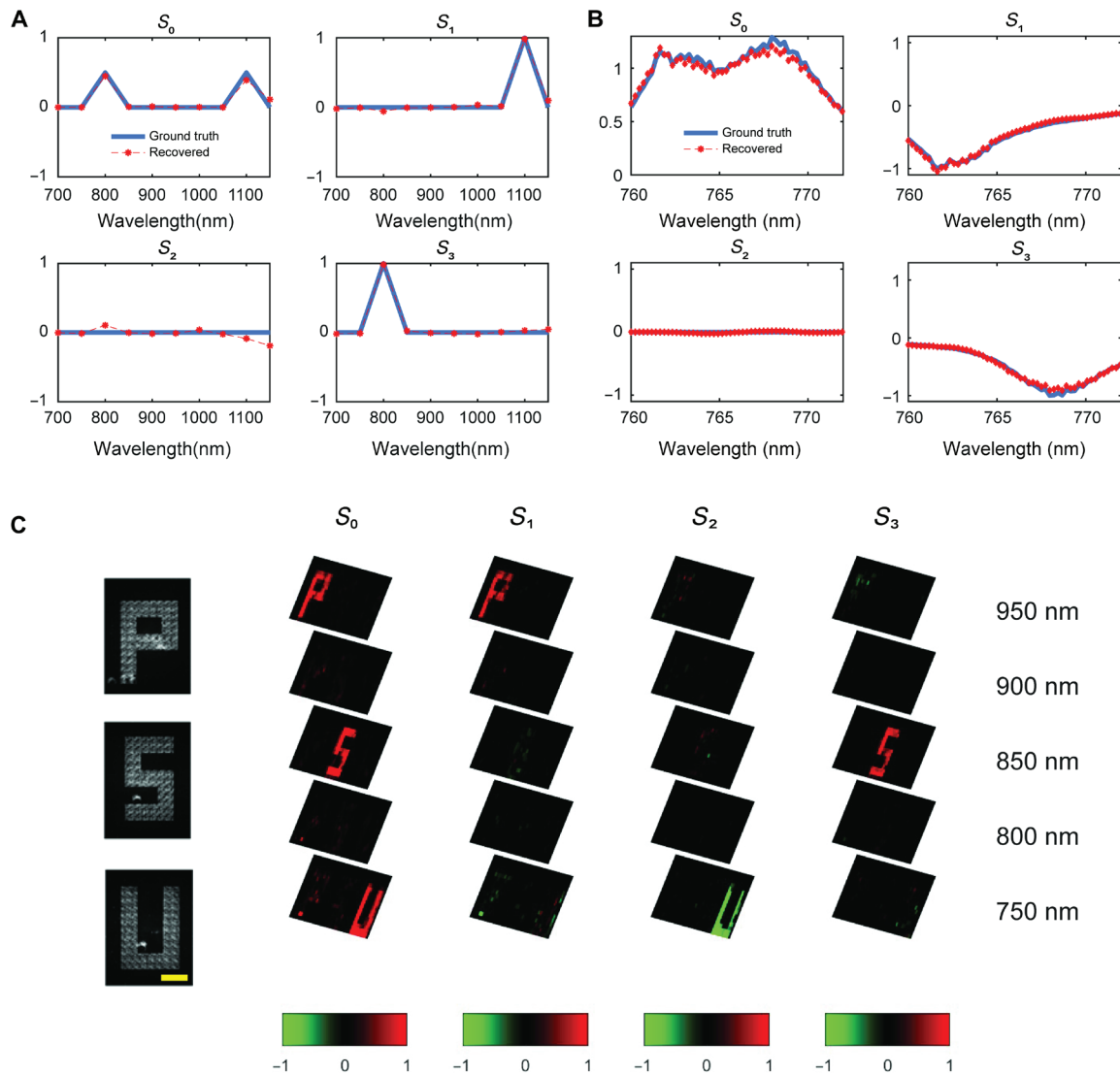


Fig. 3. Spectro-polarimetry and hyperspectro-polarimetric imaging using the optimization methods. (A) Recovery results for a broadband wavelength range (700 to 1150 nm). (B) Recovery results from 760 to 776 nm with a step size of 0.23 nm. (C) Three images on the left are the raw images captured by metasurfaces. The right figure shows HSP imaging recovery results of three letters PSU with the optimization methods. Scale bar, 600 μm . The size of the raw image is 1536 $\times$ 2048 pixels, while the size of each retrieved HSP image is 42 $\times$ 17 pixels.

right circular polarization (RCP) at 770.2 nm. The recovered Stokes spectra again showed that the recovery results matched the input states with a spectral sensitivity of 0.23 nm (Fig. 3B). The recovered spectra resolved minor oscillating peaks, demonstrating the accuracy and high spectral resolving power of our device.

To validate the imaging performance, we conducted HSP imaging with a test target. The test target consists of the letters “PSU” (see the fabrication process in Materials and Methods) with each letter being illuminated by laser beams of different polarizations and wavelengths. An inverse telescope system, comprising two lenses ($f_1 = 50$ mm and $f_2 = 150$ mm), projected the target onto the metasurface and camera (fig. S5). The captured raw images and the recovered HSP images are shown in Fig. 3C. Clearly, the three PSU letters exhibit wavelengths of 950, 850, and 750 nm, respectively. In addition, the “P” letter is in x polarization, “U” in -45° polarization,

and “S” in LCP. These results are sharp, with low background noise, and closely match the ground truth. For single image recovery, we tested both local and global optimization methods, which required 3 min and around 2 hours, respectively. To quantitatively evaluate the quality of the recovered HSP images, we calculated the structural similarity index measure (SSIM) between the results shown in Fig. 3C and their corresponding ground truth. We observed that the SSIM values (fig. S7) are almost all greater than 0.6 for all Stocks’ parameters at all wavelengths, with an average of about 0.9. Notably, at some wavelengths, the SSIM values are close to unity because the images are almost blank.

The high SSIM values indicate that the recovered HSP images capture most of the features from the testing object. However, the results in Fig. 3C still exhibit noise and artifacts. These issues likely stem from several possible sources. First, misalignment of optical

components introduced phase errors and misfocus, and power fluctuations of the light source resulted in varying illumination intensity and its distribution during the calibration and imaging experiments. Future improvements could include refining the experimental setup to ensure better alignment of optical components and stabilizing the light source. Second, our current metasurface includes only 100 different types of meta-atoms and 20×20 superpixels, which provide limited spatial resolution. This can be improved by increasing the number of meta-atoms in the metasurface and creating more superpixels. Using more superpixels could enhance image quality by providing finer spatial resolution. Furthermore, engineering more types of meta-atoms with diverse spectra-polarization responses can improve spectral and polarimetric resolution of the images. Third, imperfections in the fabrication of the metasurface, such as inaccuracies in the patterning of meta-atoms and variations in their dimensions, can lead to inconsistencies in the phase and amplitude responses, resulting in unwanted diffraction and scattering. These issues can be alleviated by further improving the fabrication process of the metasurface.

Figure 3 shows the HSP responses and imaging of the three letters PSU and is the result we retrieved using the linear regression with the least squares optimization method. This method, while effective in certain scenarios, is slow and highly susceptible to noise and errors, leading to reduced accuracy and sometimes

severe artifacts. In contrast, the NN-based approach, used for the retrieval of the rest of the images and for real-time retrieval, offers significant advantages. The NN can learn complex patterns and relationships within the data, making it more robust against noise and errors, resulting in higher-quality image reconstruction. In addition, the NN performs inference much faster than the optimization method, enabling real-time data HSP image recovery.

NN architecture, data synthesis, and training

To accelerate the recovery process, we designed a ML NN to decode the image information. The workflow of the decoding process is depicted in Fig. 4A. Once the camera captures an image, it is divided into small tiles, with each tile containing the light intensity distribution after passing through one superpixel. The spatial positions of the tiles are encoded using the one-hot method (40) (detailed in Materials and Methods). Both the tiles and their encoded positions serve as inputs for the NN. Our NN architecture comprises three hidden layers. We minimized both the number of layers and the number of neurons in each layer to achieve a compact design, thus optimizing the processing speed. The first hidden layer acts as a position-dependent system calibration layer, assigning different weights to superpixels based on their positions. The subsequent two hidden layers, serving as HSP image recovery layers, decode the spectra and polarization from each tile. The outputs are then assembled into HSP images.

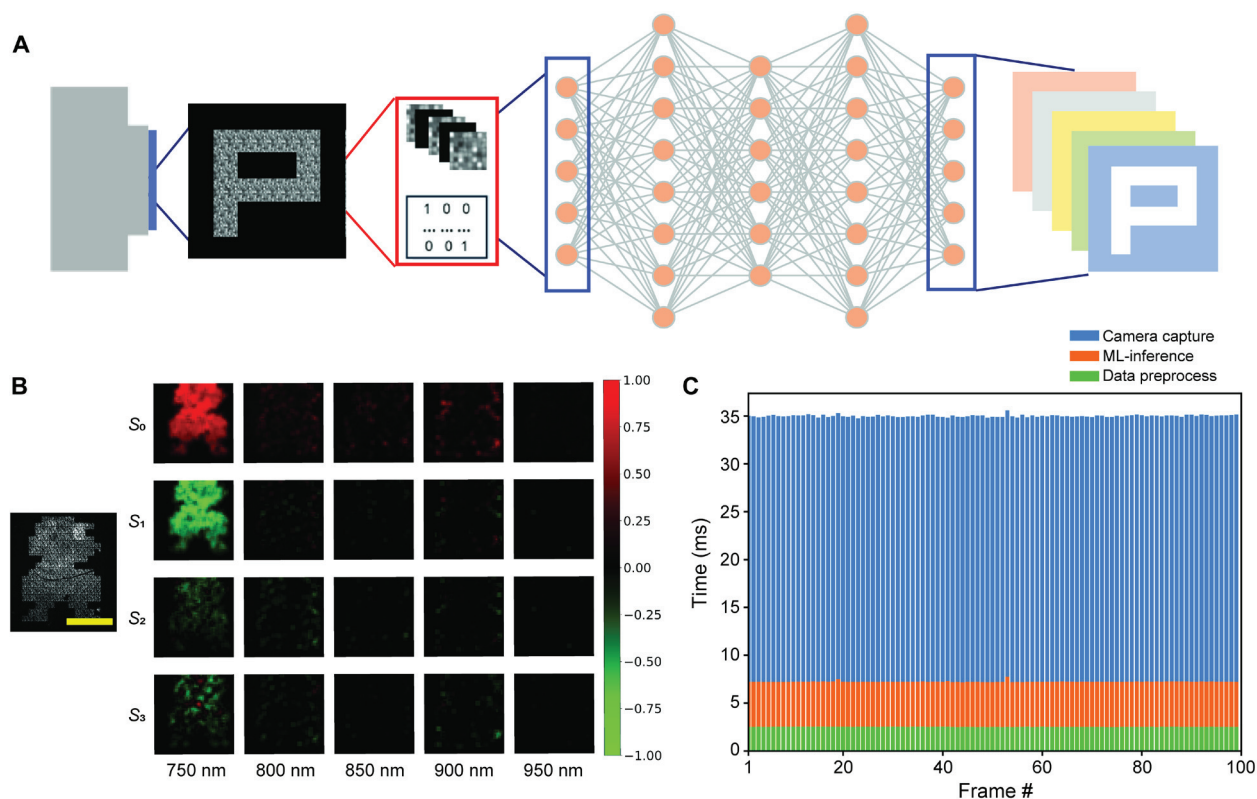


Fig. 4. ML-assisted fast recovery. (A) Architecture of the computational backend of our full-stokes HSP camera. The images captured from the HSP camera are cropped to a resolution of 1505×1764 pixels to align with the metasurface areas. Subsequently, these images are divided into segments of 42×50 superpixels. These segmented pixels, along with position encoding, are used as input for the ML model. The model, which includes three hidden layers, is responsible for position-dependent system calibration and HSP image processing. The output, consisting of 12×2100 values, is then assembled to form HSP images. (B) Reconstruction results of a cartoon figure. Horizontally polarized illumination light at 750 nm is used. The left subset shows the raw image for the cartoon figure. Scale bar, 600 μm . (C) Camera capture time, data preprocessing, and ML interface time for 100 tests.

To prepare the training dataset, our first step is to collect experimental calibration data. This was done by uniformly illuminating the metasurface with light at different wavelengths and polarization states and then capturing the corresponding output images. A total of five wavelengths spanning from 750 to 950 nm with an interval of 50 nm were used for Fig. 4 (for Fig. 5, seven wavelengths were used), and at each wavelength, 14 polarization states were prepared as the input, producing a total of 70 (5×14) calibration images. The 14 polarization states were selected to represent a uniform sampling on the surface of the polarization Poincaré sphere (see fig. S9 in the Supplementary Materials). We denote the set of the calibration data as $\{I(i, j, \lambda, P) \mid \lambda = \lambda_1, \lambda_2, \dots, \lambda_5; P = P_1, P_2, \dots, P_{14}\}$, where λ is the wavelength, P indicates the polarization state, and (i, j) represents the superpixel coordinates.

To train an NN, a large quantity of training data is required; however, experimentally generating these large training datasets is time-consuming, costly, and often not practical. Our next step is to augment the calibration data into training and validation datasets. A training image can be synthesized as follows

$$I_t(i, j) = \sum_{k=1}^5 c_k' I(i, j, \lambda_k, P_{r_k})$$

where c_k' 's are random real numbers uniformly distributed within $[0, 1.2]$ and r_k' 's are random integers selected from 1 to 14 with equal probability. In other words, through linear combinations of the calibration data, we can generate augmented training dataset with new wavelength/polarization combinations to train the NN. We also embed the positions of these superpixels (i, j) within the images into the training set in the form of one-hot encoding. This positional information can reduce errors caused by optical aberrations, metasurface fabrication defects, and other imperfections. The final size of our training set is 1.8 million. The validation dataset was constructed similarly, with the only difference being that the random weights c_k' 's are uniformly distributed within $(1.2, 1.5]$. This ensures that the validation data are different from the training data. The final size of the validation dataset is 180,000.

The training and validation datasets were then used to develop the ML model. The mean square error (MSE) was reduced to 10^{-4} upon convergence. The training time was around 12 hours using a computer equipped with an Intel i9-9900K CPU, an NVIDIA GeForce RTX 2080 Ti (11 GB) GPU, and 32 gigabyte of RAM. We validated the ML model with the synthesized validation data (fig. S10), and the recovery results matched well with the ground truth. The recovery process is significantly faster compared to the optimization-based methods, enabling real-time video recovery. During inference, processing each raw image to obtain the HSP image took less than 0.01 s, corresponding to a processing rate of over 100 FPS.

The testing dataset was prepared experimentally. We used letters (PSU), a cartoon image, and beetle specimens as testing samples. These samples were inserted into the illumination beam and placed in front of the metasurface; the response of the metasurface was then captured. For each sample, we collected experimental data images at different wavelengths and polarizations. The testing data were directly used to evaluate the trained NN model, mimicking real-world applications. We did not perform data augmentation on these experimental test images. The testing procedure is described below.

To demonstrate the ML-assisted real-time recovery, we used a cartoon figure and the PSU letters as test targets for HSP imaging (see experimental setup in fig. S5). The cartoon figure was illuminated with linearly polarized light at 750 nm. Its HSP image, recovered by our ML-assisted method from a single snapshot, shows the cartoon figure along with the wavelength and polarization information (Fig. 4B). Our method enables the acquisition of complete Stokes parameters, facilitating the accurate analysis of all polarization states, including circular or elliptical ones. This contrasts with other reported methods that can only ascertain linear polarization states (12). The noise and artifacts in Fig. 4B can be explained by the same factors as those in Fig. 3C.

Owing to the high speed of ML-assisted recovery, we demonstrated real-time HSP video acquisition. For this test, we directly fed the camera's output to our trained network and displayed the recovered HSP video in real time. During the experiment, we varied the wavelengths and polarizations of the incident light while recording. The video stream was processed by the NN and the HSP video was recovered in real-time. From the recorded HSP video (movies S1 and S2), it is evident that our HSP camera accurately captured variations in

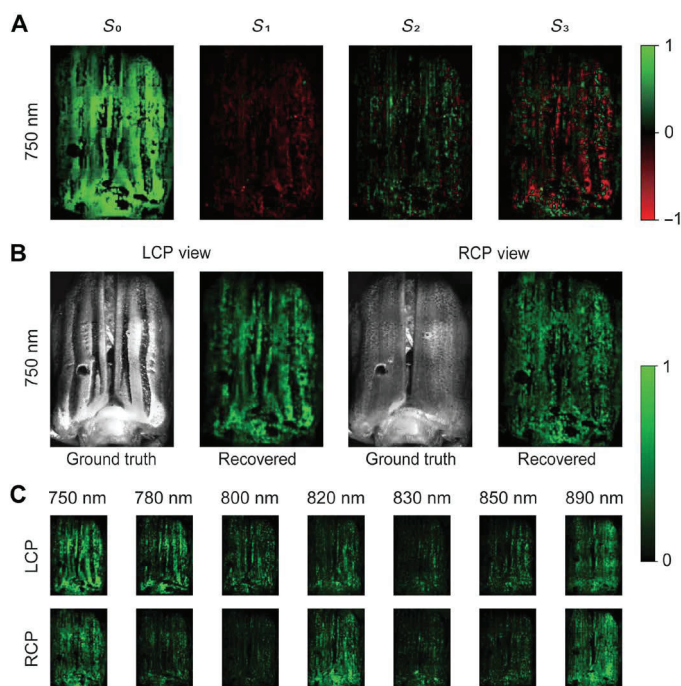


Fig. 5. Hyperspectro-polarization imaging of *Chrysina gloriosa* green gold scarab beetle. (A) Reconstruction results of *C. gloriosa* green gold scarab beetle from the ML model with laser illumination (750 nm). The training dataset includes seven wavelengths. (B) LCP and RCP views of the beetle at 750 nm: The left-side views are captured by a camera with LCP or RCP illuminations, while the right-side ones are calculated using retrieved Stokes' parameters. The RCP and LCP views are defined as the reflected image of the beetle with RCP and LCP illuminations, respectively. (C) LCP and RCP views for seven wavelengths. In the LCP views, from 750 to 850 nm, the vertical stripes on the back of the beetle can be observed. At 750 nm, the stripes are consistent with the ground truth. As the wavelength increases, the stripes on the back of the beetle gradually become less distinguishable. By 890 nm, the stripes become vague, and the LCP view is no longer significantly different from the RCP view. This is because the response of the beetle's back to polarized light changes as the wavelength increases. In the RCP view, there are no clear bright stripes on the back of the beetle; instead, there are widely distributed spots at the 750 nm, consistent with ground truth.

polarizations and wavelengths. Figure 4C breaks down the contributions from the camera readout, data preprocessing, and ML-assisted recovery to the processing time for each frame, demonstrating that our HSP camera operates at approximately 28 FPS when frames are processed synchronously. In asynchronous processing, the speed increases to 36 FPS, which corresponds to and is limited by the maximal camera readout frame rate, with a latency of 0.02 s.

To further evaluate the capability and robustness of our HSP camera, we tested it on a natural object with a more complex surface than our prior test targets. We chose a beetle specimen, *Chrysina gloriosa*, the green gold scarab beetle Arizona, known for its reflection of a high degree of LCP. We captured the light reflected by the beetle with our HSP camera (see experimental setup in fig. S12). The recovered multidimensional images (Fig. 5A) distinctly show stripes in S_3 image, indicating selective reflection of left-handed circular or elliptical polarization in those areas. From the recovered Stokes images, we calculated images of the beetle under RCP or LCP illuminations (Fig. 5B). These results closely match the experimental images captured under the same conditions. The LCP and RCP views for seven wavelengths are shown in Fig. 5C. For improving the density of the information in the HSP images, we have conducted new experiments with a glass/acrylic sample and obtained its HSP images with both 50-nm wavelength intervals and 5-nm wavelength intervals, shown in figs. S13 and S14, respectively.

Our tests have demonstrated the utility of our device across a range of test targets, including real-world objects, highlighting its robustness, accuracy, and the effectiveness of our ML approach for real-time data recovery. Our system exhibits remarkable versatility. For instance, through different calibrations, the same device can cater to diverse requirements—ranging from broad bandwidth applications with large wavelength steps (from 700 to 1150 nm) to high spectral resolution (0.23 nm) within a more limited bandwidth. Furthermore, we have the flexibility to adjust the emphasis on spatial, spectral, and polarimetric channels by allocating a varying number of metasurface encoder pixels to resolve the information within those channels. For example, we can enhance spectral and polarization resolutions at the expense of spatial resolution by incorporating more metasurface encoder pixels within each superpixel. Alternatively, we might choose to sacrifice polarization data to boost spectral performance, dedicating all pixels to constructing spectral responses within the transmission matrix. Conversely, polarization accuracy can be improved by narrowing the range of wavelengths resolved. Thus, depending on specific application scenarios, we can tailor the system's capabilities to provide on-demand multidimensional imaging.

DISCUSSION

In this work, we transformed a conventional camera into a compact HSP camera by integrating a judiciously designed metasurface. This metasurface, consisting of meta-atoms having rich polarization-dependent resonances, encodes spectral and polarimetric information into distinct intensity distributions. These distributions can be decoded using either an optimization-based or a ML-based method. Our experimental results showed that this metasurface-enhanced camera can resolve full-Stokes polarization across a broad spectral range (700 to 1150 nm) from a single snapshot, achieving a spectral sensitivity as high as 0.23 nm. The HSP camera, backed with our ML-based recovery, achieves real-time recovery on a standard laptop.

Our system records full-Stokes HSP videos in real time at 28 FPS, limited only by the camera's readout rate. Our metasurface-based camera offers a compact, speedy, and cost-effective solution for a multidimensional imaging system that harnesses spatial, spectral, and polarization information within light fields. In addition, the system's compactness allows for potential integration into other platforms, such as microfluidic systems or fiber-optic probes, which could pave the way for new applications in real-time monitoring or in situ analysis.

MATERIALS AND METHODS

Fabrication of the metasurfaces and imaging targets

We fabricated our encoding metasurfaces on indium tin oxide glass, which was thoroughly cleaned through sonication in acetone and isopropyl alcohol (IPA), each for a duration of 3 min. Following this, we grew a 600-nm layer of amorphous silicon using plasma-enhanced chemical vapor deposition. We then spun a 1:1 diluted ZEP 520A electron beam resist at 3000 rpm for 45 s and prebaked it at 180°C for 3 min. The meta-atoms were inscribed using a Vistec 5200 100 kV, followed by a 3-min development in *n*-amyl acetate and a 1-min rinse in MIBK:IPA. We then deposited 40-nm Cr films at a rate of 2 Å/s using a Temescal electron beam evaporation system. The pattern was lifted off in a water bath at 80°C for 2 hours using 1165 remover (MicroChem). To eliminate any potential contamination from residual metal particles, we sonicated the sample for several minutes in solvent before drying it with an N_2 gun. We used a chlorine-based plasma inductively coupled plasma reactive-ion etching (ICP-RIE) recipe, involving Cl_2 and Ar gas, to etch the amorphous Si and create the meta-atoms. Last, we immersed the sample in a Chromium etchant to remove the mask.

We fabricated the imaging targets, featuring PSU letters and a cartoon figure, on a glass substrate. The substrate was thoroughly cleaned through sonication in acetone and IPA, each for 3 min. We then deposited a 150-nm Al layer at a rate of 2 Å/s using a Temescal electron beam evaporation system. Subsequently, we spun the SPR950 photoresist at 3000 rpm for 45 s and prebaked it at 90°C for 1 min. The patterns were written using a Heidelberg MLA 150 Direct Write Exposure Tool, followed by a 1-min development in CD26 and a 1-min water rinse. We then dry-etched the Al with a resist mask using a chlorine-based plasma ICP-RIE recipe, which involved Cl_2 and BCl_3 gas. Last, we immersed the sample in Remover PG to remove the resist.

Optical characterization

The characterization setup for the encoding metasurfaces consists of three parts: a tunable source, a beam expander, and an HSP imaging system, as shown in fig. S3. The tunable source comprises a Ti:sapphire-pulsed laser, capable of tuning the wavelength across a broad range, and a half waveplate and quarter waveplate for full-Stokes polarization tuning. We used two singlet lenses ($f_1 = 50$ mm and $f_2 = 150$ mm) to form a simple microscope setup, providing $\times 3$. The role of this microscope setup is twofold. In the calibration stage, it expands the input light beam and ensures that the resulting beam size is sufficiently large to illuminate the entire area of our metasurface. In the measurement stage, it projects the image of the target onto the metasurface. Essentially, the metasurface can be directly integrated on the top of the camera sensor. However, in our current studies, to maintain the flexibility to iterate the metasurface design

without damaging the camera each time, a microscope system was used after the metasurface to “virtually” attach the metasurface onto the camera sensor (fig. S3).

To calibrate the system, we varied the input wavelength and the polarization state using the tunable light source and then captured the resultant images using the HSP imaging system. An example of the calibration results is depicted in Fig. 2D. Following the same procedure, we used illumination light with different wavelength and polarization state combinations to test our HSP imaging system. The captured raw images were used to retrieve the illumination spectra and polarization states (Fig. 3, A and B) using the optimization method.

For acquiring HSP images, the target objects were placed at the front focal plane of the first singlet lens ($f_1 = 50$ mm). We used PSU letters and a cartoon figure as the test targets. The raw images captured by the HSP imaging system were then used to reconstruct the HSP images (Fig. 3C). For acquiring the HSP images of the beetle, we modified the setup to measure the reflected light from the object instead of the transmitted light (fig. S6): The light source was coupled into a single-mode fiber and passed through a diffuser to ensure the light beam had uniform intensity distribution. Next, it passed through a linear polarizer and became x (horizontally)-polarized. The beam illuminated the beetle sample, which was placed at the front focal plane of the first singlet lens. The reflected light from the beetle was lastly collected by the HSP imaging system.

Data preprocessing in ML process

The raw image taken by the camera is of the size of 1505×1764 , which is relatively large for the input to the ML model. To simplify the model for high-speed imaging and future potential mobile deployment, we divide the image into superpixels, and one superpixel is 88×88 . We then resize the image to 44×44 to reduce the image size and to make sure we still have enough superpixels to cover the intensity information for wavelength and polarization retrieval. The 88×88 sliding window shifts 44 pixels to make another superpixel, and in total, we have 42×50 image pixels from one shot (fig. S8). All camera pixels within the superpixel are flattened to 1D. Because we have 42×50 image pixels in total, there are 2100 unique image pixels. Therefore, we have 4036 numerical values as the input including the images pixel information and position information, and the output is the spectrum and polarization information for the image pixel.

Supplementary Materials

The PDF file includes:

Notes S1 to S6

Figs. S1 to S14

Legends for movies S1 and S2

Other Supplementary Material for this manuscript includes the following:

Movies S1 and S2

REFERENCES AND NOTES

- J. Park, X. Feng, R. Liang, L. Gao, Snapshot multidimensional photography through active optical mapping. *Nat. Commun.* **11**, 5602 (2020).
- S. Liu, X. Liu, J. Yuan, J. Bao, Multidimensional information encryption and storage: When the input is light. *Research* **2021**, 7897849 (2021).
- G. Cai, Y. Li, Y. Zhang, X. Jiang, Y. Chen, G. Qu, X. Zhang, S. Xiao, J. Han, S. Yu, Y. Kivshar, Q. Song, Compact angle-resolved metasurface spectrometer. *Nat. Mater.* **23**, 71–78 (2024).
- Z. Shen, F. Zhao, C. Jin, S. Wang, L. Cao, Y. Yang, Monocular metasurface camera for passive single-shot 4D imaging. *Nat. Commun.* **14**, 1035 (2023).
- G. Lu, B. Fei, Medical hyperspectral imaging: A Review. *J. Biomed. Opt.* **19**, 010901 (2014).
- A. Khan, A. D. Vibhute, S. Mali, C. H. Patil, A systematic review on hyperspectral imaging technology with a machine and deep learning methodology for agricultural applications. *Eco. Inform.* **69**, 101678 (2022).
- C. He, H. He, J. Chang, B. Chen, H. Ma, M. J. Booth, Polarisation optics for biomedical and clinical applications: A review. *Light Sci. Appl.* **10**, 194 (2021).
- M. F. Sterzik, S. Bagnulo, E. Palle, Biosignatures as revealed by spectropolarimetry of earthshine. *Nature* **483**, 64–66 (2012).
- X. Zhou, L. Ma, W. Brown, J. V. Little, A. Y. Chen, L. L. Myers, B. D. Sumer, B. Fei, Automatic detection of head and neck squamous cell carcinoma on pathologic slides using polarized hyperspectral imaging and machine learning. *Proc. SPIE Int. Soc. Opt. Eng.* **11603**, 116030Q (2021).
- D. Tua, R. Liu, W. Yang, L. Zhou, H. Song, L. Ying, Q. Gan, Imaging-based intelligent spectrometer on a plasmonic rainbow chip. *Nat. Commun.* **14**, 1902 (2023).
- J. Zuo, J. Bai, S. Choi, A. Basiri, X. Chen, C. Wang, Y. Yao, Chip-integrated metasurface full-stokes polarimetric imaging sensor. *Light Sci. Appl.* **12**, 218 (2023).
- A. Altaqui, P. Sen, H. Schrickx, J. Rech, J.-W. Lee, M. Escuti, W. You, B. J. Kim, R. Kolbas, B. T. O'Connor, M. Kudenov, Mantis Shrimp-inspired organic photodetector for simultaneous hyperspectral and polarimetric imaging. *Sci. Adv.* **7**, eabe3196 (2021).
- X. Tu, O. J. Spire, X. Tian, N. Brock, R. Vail, V. Gruet, Division of amplitude RGB full-stokes camera using micro-polarizer arrays. *Opt. Express* **25**, 33160–33175 (2017).
- K. Shinoda, Y. Ohtera, M. Hasegawa, Snapshot multispectral polarization imaging using a photonic crystal filter array. *Opt. Express* **26**, 15948–15961 (2018).
- M. Garcia, C. Edmiston, R. Marinov, A. Vail, V. Gruet, Bio-inspired color-polarization imager for real-time in situ imaging. *Optica* **4**, 1263 (2017).
- S. Sattar, P.-J. Lapray, L. Aksas, A. Foulonneau, L. Bigué, Snapshot spectropolarimetric imaging using a pair of filter array cameras. *Opt. Express* **61**, 043104 (2022).
- S. Sattar, P.-J. Lapray, A. Foulonneau, L. Bigué, Review of spectral and polarization imaging systems, in *Unconventional Optical Imaging II*, C. Fournier, M. P. Georges, G. Popescu, Eds. (International Society for Optics and Photonics, 2020).
- N. Yu, P. Genevet, M. A. Kats, F. Aieta, J.-P. Tetienne, F. Capasso, Z. Gaburro, Light propagation with phase discontinuities: Generalized laws of reflection and refraction. *Science* **334**, 333–337 (2011).
- X. Ni, N. K. Emani, A. V. Kildishev, A. Boltasseva, V. M. Shalae, Broadband light bending with plasmonic nanoantennas. *Science* **335**, 427 (2012).
- F. Ding, A. Pors, S. I. Bozhevolnyi, Gradient Metasurfaces: A review of fundamentals and applications. *Rep. Prog. Phys.* **81**, 026401 (2018).
- S. Chang, X. Guo, X. Ni, Optical metasurfaces: progress and applications. *Annu. Rev. Mat. Res.* **48**, 279–302 (2018).
- X. Ni, Z. J. Wong, M. Mrejen, Y. Wang, X. Zhang, An ultrathin invisibility skin cloak for visible light. *Science* **349**, 1310–1314 (2015).
- T. Sun, J. Hu, X. Zhu, F. Xu, C. Wang, Broadband single-chip full Stokes polarization-spectral imaging based on all-dielectric spatial multiplexing Metasens. *Laser Photonics Rev.* **16**, 2100650 (2022).
- X. Wang, Z. Yang, F. Bao, T. Sentz, Z. Jacob, Spinning Meta-Cam: spectro-polarimetric long-wave infrared thermal imaging based on spinning metasurfaces. arXiv:2301.11519 [physics.optics] (23 September 2023).
- Z. Wang, S. Yi, A. Chen, M. Zhou, T. S. Luk, A. James, J. Nogan, W. Ross, G. Joe, A. Shahsafi, K. X. Wang, M. A. Kats, Z. Yu, Single-shot on-chip spectral sensors based on photonic crystal slabs. *Nat. Commun.* **10**, 1020 (2019).
- B. Redding, S. F. Liew, R. Sarma, H. Cao, Compact spectrometer based on a disordered photonic chip. *Nat. Photon.* **7**, 746–751 (2013).
- Z. Yang, T. Albrow-Owen, H. Cui, J. Alexander-Webber, F. Gu, X. Wang, T.-C. Wu, M. Zhuge, C. Williams, P. Wang, A. V. Zayats, W. Cai, L. Dai, S. Hofmann, M. Overend, L. Tong, Q. Yang, Z. Sun, T. Hasan, Single-nanowire spectrometers. *Science* **365**, 1017–1020 (2019).
- J. Xiong, X. Cai, K. Cui, K. Cui, Y. Huang, Y. Huang, J. Yang, H. Zhu, W. Li, B. Hong, B. Hong, S. Rao, Z. Zheng, S. Xu, Y. He, F. Liu, F. Liu, X. Feng, W. Zhang, W. Zhang, Dynamic brain spectrum acquired by a real-time ultraspectral imaging chip with reconfigurable metasurfaces. *Optica* **9**, 461–468 (2022).
- J. Yang, K. Cui, X. Cai, J. Xiong, H. Zhu, S. Rao, S. Xu, Y. Huang, F. Liu, X. Feng, W. Zhang, Ultraspectral imaging based on metasurfaces with freeform shaped meta-atoms. *Laser Photonics Rev.* **16**, 2100663 (2022).
- C.-C. Chang, H.-N. Lee, On the estimation of target spectrum for filter-array based spectrometers. *Opt. Express* **16**, 1056–1061 (2008).
- Y. Ni, C. Chen, S. Wen, X. Xue, L. Sun, Y. Yang, Computational spectropolarimetry with a tunable liquid crystal metasurface. *eLight* **2**, 23 (2022).
- W. T. Chen, P. Török, M. R. Foreman, C. Y. Liao, W.-Y. Tsai, P. R. Wu, D. P. Tsai, Integrated plasmonic metasurfaces for spectropolarimetry. *Nanotechnology* **27**, 224002 (2016).
- S. Zhang, Y. Dong, H. Fu, S.-L. Huang, L. Zhang, A spectral reconstruction algorithm of miniature spectrometer based on sparse optimization and dictionary learning. *Sensors* **18**, 644 (2018).

34. S. P. Boyd, L. Vandenberghe *Convex Optimization*. (Cambridge University Press, Cambridge, 2013).
35. M. J. D. Powell A fast algorithm for nonlinearly constrained optimization calculations. In *Numerical Analysis*; Watson, G. A., Ed. (Springer, Berlin, 1978).
36. S. P. Han, A globally convergent method for nonlinear programming. *J. Optim. Theory Appl.* **22**, 297–309 (1977).
37. M. R. Bonyadi, Z. Michalewicz, Particle Swarm optimization for single objective continuous space problems: A review. *Evol. Comput.* **25**, 1–54 (2017).
38. S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trends Mach. Learn.* **3**, 1–122 (2011).
39. L. Kang, C.-Y. Wang, X. Guo, X. Ni, Z. Liu, D. H. Werner, Nonlinear chiral meta-mirrors: Enabling technology for ultrafast switching of light polarization. *Nano Lett.* **20**, 2047–2055 (2020).
40. J. T. Hancock, T. M. Khoshgoftaar, Survey on categorical data for neural networks. *J. Big Data* **7**, 28 (2020).

Acknowledgment

Funding: The work was partially supported by the National Science Foundation under grant agreements ECCS-2047446, ECCS-2114266, and CHE-2305139 and the National Eye Institute, National Institutes of Health under grant agreement 1R21EY031853-01.

Author contributions: Conceptualization: L.Z. and X.N. Methodology: L.Z., C.Z., Y.D., H.-J.A., S.C., Y.M.D., B.L., M.T.R., T.X., X.C., Z.L., and X.N. Investigation: L.Z., C.Z., B.L., Z.L., and X.N. Visualization: L.Z., C.Z., B.L., Z.L., and X.N. Supervision: X.N. and Z.L. Writing—original draft: L.Z. and C.Z. Writing—review and editing: L.Z., C.Z., Y.D., H.-J.A., S.C., Y.M.D., B.L., M.T.R., Z.L., and X.N.

Competing interests: The other 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 28 March 2024

Accepted 29 July 2024

Published 4 September 2024

10.1126/sciadv.adp5192