Check for updates

REVIEW SUMMARY

OPTICAL SENSING

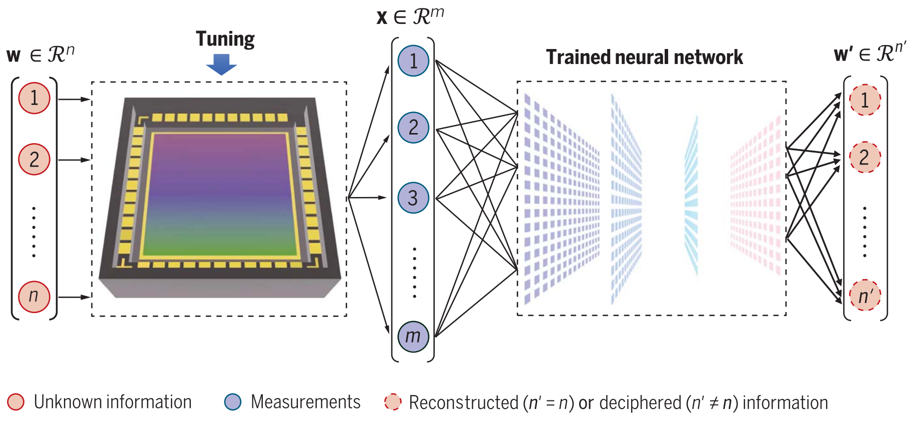
Geometric deep optical sensing

Shaofan Yuan†, Chao Ma†, Ethan Fetaya, Thomas Mueller*, Doron Naveh*, Fan Zhang*, Fengnian Xia*

BACKGROUND: Optical sensing devices measure the rich physical properties of an incident light beam, such as its power, polarization state, spectrum, and intensity distribution. Most conventional sensors, such as power meters, polarimeters, spectrometers, and cameras, are monofunctional and bulky. For example, classical Fourier-transform infrared spectrometers and polarimeters, which characterize the optical spectrum in the infrared and the polarization state of light, respectively, can occupy a considerable portion of an optical table. Over the past decade, the development of integrated sensing solutions by using miniaturized devices together with advanced machine-learning algorithms has accelerated rapidly, and optical sensing research has evolved into a highly interdisciplinary field that encompasses devices and materials engineering, condensed matter physics, and machine learning. To this end, future optical sensing technologies will benefit from innovations in device architecture, discoveries of new quantum materials, demonstrations of previously uncharacterized optical and optoelectronic phenomena, and rapid advances in the development of tailored machine-learning algorithms.

ADVANCES: Recently, a number of sensing and imaging demonstrations have emerged that differ substantially from conventional sensing schemes in the way that optical information is detected. A typical example is computational spectroscopy. In this new paradigm, a compact spectrometer first collectively captures the comprehensive spectral information of an incident light beam using multiple elements or a single element under different operational states and generates a high-dimensional photoresponse vector. An advanced algorithm then interprets the vector to achieve reconstruction of the spectrum. This scheme shifts the physical complexity of conventional grating- or interference-based spectrometers to computation. Moreover, many of the recent developments go well beyond optical spectroscopy, and we discuss them within a common framework, dubbed “geometric deep optical sensing.” The term “geometric” is intended to emphasize that in this sensing scheme, the physical properties of an unknown light beam and the corresponding photoresponses can be regarded as points in two respective high-dimensional vector spaces and that the sensing process can be considered to be a mapping from one vector space to the other. The mapping can be linear, nonlinear, or highly entangled; for the latter two cases, deep artificial neural networks represent a natural choice for the encoding and/or decoding processes, from which the term “deep” is derived. In addition to this classical geom view, the quantum geometry of Bloch electrons in Hilbert space, such as Berry curvature and quantum metrics, is essential for the determination of the polarization-dependent photoresponses in some optical sensors. In this Review, we first present a general perspective of this sensing scheme from the viewpoint of information theory, in which the photoresponse measurement and the extraction of light properties are deemed as information-encoding and -decoding processes, respectively. We then discuss demonstrations in which a reconfigurable sensor (or an array thereof), enabled by device reconfigurability and the implementation of neural networks, can detect the power, polarization state, wavelength, and spatial features of an incident light beam.

OUTLOOK: As increasingly more computing resources become available, optical sensing is becoming more computational, with device reconfigurability playing a key role. On the one hand, advanced algorithms, including deep neural networks, will enable effective decoding of high-dimensional photoresponse vectors, which reduces the physical complexity of sensors. Therefore, it will be important to integrate memory cells near or within sensors to enable efficient processing and interpretation of a large amount of photoresponse data. On the other hand, analog computation based on neural networks can be performed with an array of reconfigurable devices, which enables direct multiplexing of sensing and computing functions. We anticipate that these two directions will become the engineering frontier of future deep sensing research. On the scientific frontier, exploring quantum geometric and topological properties of new quantum materials in both linear and nonlinear light-matter interactions will enrich the information-encoding pathways for deep optical sensing. In addition, deep sensing schemes will continue to benefit from the latest developments in machine learning. Future highly compact, multifunctional, reconfigurable, and intelligent sensors and imagers will find applications in medical imaging, environmental monitoring, infrared astronomy, and many other areas of our daily lives, especially in the mobile domain and the internet of things. ■



Schematic of deep optical sensing. The n-dimensional unknown information ($\mathbf{w}$) is encoded into an m-dimensional photoresponse vector ($\mathbf{x}$) by a reconfigurable sensor (or an array thereof), from which $\mathbf{w'}$ is reconstructed by a trained neural network ($n' = n$ and $\mathbf{w'} \approx \mathbf{w}$). Alternatively, $\mathbf{x}$ may be directly deciphered to capture certain properties of $\mathbf{w}$. Here, $\mathbf{w}$, $\mathbf{x}$, and $\mathbf{w'}$ can be regarded as points in their respective high-dimensional vector spaces $\mathcal{R}^n$, $\mathcal{R}^m$, and $\mathcal{R}^{n'}$.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: thomas.mueller@tuwien.ac.at (T.M.); doron.naveh@biu.ac.il (D.N.); zhang@utdallas.edu (F.Z.); fengnian.xia@yale.edu (F.X.)
†These authors contributed equally to this work.

Cite this article as S. Yuan *et al.*, *Science* **379**, eade1220 (2023). DOI: 10.1126/science.ade1220

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.ade1220

Downloaded from https://www.science.org at University of Texas Dallas on May 11, 2023

REVIEW

OPTICAL SENSING

Geometric deep optical sensing

Shaofan Yuan[1]†, Chao Ma[1]†, Ethan Fetaya[2], Thomas Mueller[3]*, Doron Naveh[2]*, Fan Zhang[4,5]*, Fengnian Xia[1]*

Geometry, an ancient yet vibrant branch of mathematics, has important and far-reaching impacts on various disciplines such as art, science, and engineering. Here, we introduce an emerging concept dubbed "geometric deep optical sensing" that is based on a number of recent demonstrations in advanced optical sensing and imaging, in which a reconfigurable sensor (or an array thereof) can directly decipher the rich information of an unknown incident light beam, including its intensity, spectrum, polarization, spatial features, and possibly angular momentum. We present the physical, mathematical, and engineering foundations of this concept, with particular emphases on the roles of classical and quantum geometry and deep neural networks. Furthermore, we discuss the new opportunities that this emerging scheme can enable and the challenges associated with future developments.

Light sensors are ubiquitous and essential in many aspects of our lives. In humans, it is believed that more than 80% of the total information captured by the five senses is perceived by the eyes (1)—light sensors in the visible spectral range—as a result of evolution and natural selection over millions of years. There are also many different types of human-made light sensors, and such a sensor is usually built to probe a specific physical property of light. For example, an imager generates a two-dimensional (2D) map of light intensity, a spectrometer determines the spectral composition of light, and a polarimeter measures the polarization state of light. Many conventional light sensors are bulky, expensive, and monofunctional. In the past decade, as sensing tasks have become more demanding and as more computational resources have become available, two trends have emerged in optical sensing. First, it has become critical to build miniaturized, inexpensive sensors that can be integrated on-chip to enable pervasive applications, especially in mobile domains such as mobile phones, smart watches, autonomous vehicles, robots, and drones. This trend is evidenced by the demonstrations of ultracompact spectrometers in various spectral ranges (2, 3) that use miniaturized dispersive optical components (4–8), on-chip interferometers (9–12), arrays of sensors with different spectral responses (13–24), or spectrally reconfigurable sensors (25–28). Moreover, on-chip polarization detectors (29–31) and compact spectral imagers (32–34) have also been extensively investigated. Second, algorithms are playing increasingly important roles in sensing, and many recent developments have leveraged machine-learning algorithms such as regression techniques and neural networks in sensor design and operation (2, 3, 35, 36).

Here, in addition to covering some miniaturized spectral sensors, we review several innovative optical sensing schemes in which the functions of a miniaturized sensor go beyond those of traditional concepts (37–47). The recent progress in this field has been enabled by innovations in device physics and the implementation of advanced machine-learning algorithms. We approach these schemes within a common framework that we call "geometric deep optical sensing" [not to be confused with "geometric deep learning," a field that seeks to understand neural networks in non-Euclidean domains (48)]. The term "geometric" is intended to emphasize that the physical properties of the unknown light and the corresponding photoresponse can be regarded as points in two respective high-dimensional vector spaces and that the sensing process can be regarded as a mapping from one vector space to the other. The mapping can be linear, nonlinear, or highly entangled; for the latter two cases, deep artificial neural networks represent a natural choice for the encoding and/or decoding processes (49), from which the term "deep" is derived. In addition to the geometric perspective discussed above, the quantum geometry of Bloch electrons in Hilbert space, such as Berry curvature and quantum metrics, plays an important role in generating the polarization-dependent photoresponse vectors in some of the demonstrations (47, 50–60).

An information theory view

In general, from an information theory (61) perspective, an (optical) sensing process can be understood as follows (Fig. 1A): A sensor acts as an encoder that converts unknown, high-dimensional physical quantities into sensor outputs; a channel corresponding to a noisy measurement process reads the sensor outputs; and a decoder deciphers the encoded high-dimensional information. Here, the high-dimensional physical quantities can be characterized by a vector $\mathbf{w}$, which represents the intrinsic physical properties of a light beam, such as power, spectrum, polarization state, spatial or temporal properties, or the combination of several of these. The vector $\mathbf{w}$ can be treated as a point in a vector space of dimension n ($\mathbf{w} \in \mathcal{R}^n$) (Fig. 1B). In traditional sensing schemes, direct determination of such a vector requires a series of measurements that use different types of optical components such as beam splitters, waveplates, filters, dispersive gratings, and power meters, followed by data processing steps. In the sensing scheme introduced here, $\mathbf{w}$ is first encoded into a response vector $\mathbf{x}$ by a single sensor or an array thereof, which is engineered to capture spatial, spectral, polarization, and/or temporal information. Vector $\mathbf{x}$ can be treated as a point in a vector space of dimension m ($\mathbf{x} \in \mathcal{R}^m$) (Fig. 1B). It may be interpreted directly to capture certain properties of $\mathbf{w}$ or decoded into a vector $\mathbf{w'}$ to reconstruct the desired physical quantities of $\mathbf{w}$.

In contrast to traditional sensing schemes, in the geometric deep optical sensing scheme, both the encoding and decoding processes can be implicit. Moreover, different kinds of information can be encoded concurrently into the sensor outputs. As a result, this sensing scheme allows for the detection of multiple physical properties of light and functionality multiplexing. Depending on the dimensionality of the two vector spaces $\mathcal{R}^n$ and $\mathcal{R}^m$, we distinguish between three cases. The case $m = n$ is the most common and corresponds, for example, to the case of computational spectrometers (13, 18, 20, 25). The case $m < n$ corresponds to compressed sensing (62–65). By using prior knowledge or proper assumptions, high-dimensional information can be reconstructed from a low-dimensional photoresponse. Finally, the case $m > n$ may have the advantage of being more robust to noise because of redundancy introduced in the information-encoding process (61).

Information-encoding mechanisms

As shown in Fig. 1B, the encoder's role is to map the physical information in $\mathcal{R}^n$ to a photoresponse in $\mathcal{R}^m$. In the following sections, we

---

[1]Department of Electrical Engineering, Yale University, New Haven, CT, USA. [2]Faculty of Engineering, Bar-Ilan University, Ramat Gan, Israel. [3]Institute of Photonics, Vienna University of Technology, Vienna, Austria. [4]Department of Physics, The University of Texas at Dallas, Richardson, TX, USA. [5]Department of Physics, Massachusetts Institute of Technology, Cambridge, MA, USA.
*Corresponding author. Email: thomas.mueller@tuwien.ac.at (T.M.); doron.naveh@biu.ac.il (D.N.); zhang@utdallas.edu (F.Z.); fengnian.xia@yale.edu (F.X.)
†These authors contributed equally to this work.

Downloaded from https://www.science.org at University of Texas Dallas on May 11, 2023

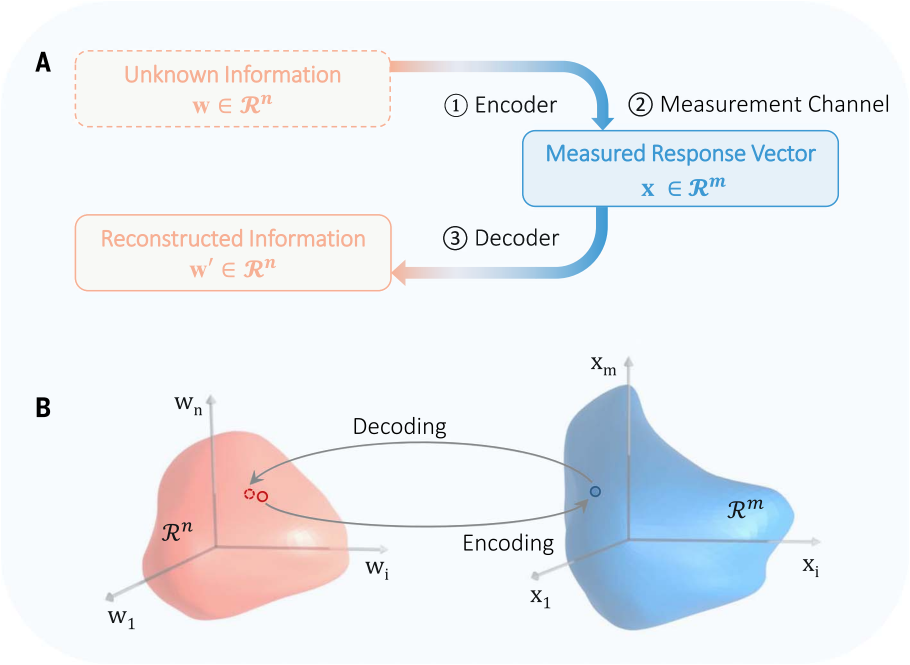


Fig. 1. An information theory view of deep optical sensing. (**A**) An encoder (an optical sensor or sensor array) converts the unknown *n*-dimensional physical information $\mathbf{w} \in \mathcal{R}^n$ into electrical outputs, the channel corresponds to a noisy measurement process that reads the *m*-dimensional output $\mathbf{x} \in \mathcal{R}^m$, and a decoder reconstructs the information $\mathbf{w}' \in \mathcal{R}^n$. (**B**) The vectors **w** and **x** can be regarded as points in *n*- and *m*-dimensional vector spaces $\mathcal{R}^n$ and $\mathcal{R}^m$, respectively. A mathematical tool, for example, a trained neural network, maps **x** to **w′** from $\mathcal{R}^m$ back to $\mathcal{R}^n$. A high-performance sensor captures the unknown information accurately such that the reconstructed **w′** is close to **w** in $\mathcal{R}^n$. **w** and **w′** are represented by the red hollow dots with solid and dotted edges, respectively, in $\mathcal{R}^n$. **x** is represented by the blue hollow dot in $\mathcal{R}^m$. Alternatively, **x** can be evaluated directly to capture certain features of **w**.

discuss examples of how this can be accomplished for different properties of incident light.

Tuning device geometric features and quantum geometry for polarization encoding

Geometry plays a critical role in light-matter interactions. For example, optical devices with different geometric features can exhibit distinct polarization-dependent responses to light. Previously, polarization imaging has been demonstrated by using a 2D grating matrix (*66*) consisting of optical elements with different geometries (Fig. 2A), in which different polarization components of an incident light beam are separated spatially for the subsequent polarization information–encoding process.

Other than the geometric features of the sensor, the quantum geometry of Bloch electrons, that is, Berry curvature and quantum metric, can also be tuned for the encoding of polarization information. Quantum geometry represents the geometry of the quantum states in Hilbert space, and it is critical for nonlinear photoresponses such as the second-harmonic generation and the bulk photovoltaic effect (BPVE). One notable feature of the BPVE is its strong polarization and wavelength dependencies. In the device reported by Ma *et al.* (*47*), the active material is twisted double-bilayer graphene (TDBG) sandwiched between two hexagonal boron nitride (hBN) thin films. The graphene (top) and silicon (bottom) gate electrodes are used to produce two electric potentials through hBN to modulate the quantum geometric properties of TDBG for polarization encoding.

Under excitation with linearly polarized light, the BPVE (shift current) is determined by two independent conductivity elements σ_{xxx} and σ_{yyy} that can be directly calculated by integrating $S_{xxx(yyy)}$, which is the contribution from an electron-hole pair that participates in the resonant optical transition (*51*–*54*). Here, $S_{xxx(yyy)}$ depends on the Fermi distribution difference between the electron and hole Bloch bands, the interband non-Abelian Berry connections, and the excitation light frequency. The integrand S_{xxx} in the moiré Brillouin zone is shown in Fig. 2B for Fermi energy $E_F = 0$ and interlayer potential difference $\Delta V = 100$ meV under excitation with 7.7-μm light in TDBG (*47*). Hotspots in Fig. 2B indicate the positions in momentum space at which the quantum geometric properties relevant to σ_{xxx} are pronounced, which allows for the resonant optical transition. Under circularly polarized light excitation, the BPVE (injection current) is governed by the interband Berry curvature dipoles or the Hermitian metrics (*54*).

Importantly, not only does the nonlinear response depend on the polarization of the light, but this dependence is also tunable. The tunability originates from the fact that the two gate voltages can independently control the Fermi energy (or carrier density) and the out-of-plane displacement field of the 2D material in the device. The displacement field can modulate the band structure that specifically determines the quantum geometric properties of Bloch states. The Fermi energy determines the electron-hole pairs of Bloch states that are available for the resonant optical transition. Together, they determine the nonlinear conductivity tensor and hence tune the BPVE. As a result, the polarization information of the incident light can be encoded into the nonlinear photoresponse map that is generated under different pairs of biasing gate voltages (*47*). Such an encoding process is implicit because of the complexity of the device, including strain, disorder, inhomogeneity, and so on (*67*). However, the decoding can still be successfully performed by using a trained neural network, as discussed below.

Engineering the spectral response for optical spectroscopy

Optical elements with different geometric features can also be directly integrated with complementary metal-oxide semiconductor (CMOS) sensors to encode the spectral information (Fig. 2C) (*19*). In such a spectrometer, each element captures certain spectral characteristics by leveraging the rationally designed geometric features. Indeed, this approach has been extensively used for information encoding in spectral and polarization imaging (*17*, *19*, *68*–*70*). However, despite the effectiveness of such a geometric approach, the physical layout of optical elements can hardly be reconfigured after fabrication, which limits their potential in advanced applications. For example, to achieve high resolution in spectroscopy, a large number of elements with different geometric features are required, yet their scaling is limited by optical diffraction, which results in a large overall device footprint. Moreover, because of the lack of reconfigurability, it is difficult to fully leverage the capacities of machine-learning algorithms for deciphering nontrivial high-dimensional data.

In addition to using elements with different geometric features, there are a number of other approaches for encoding spectral information for optical spectroscopy. A prime example is the engineering of the bandgap of semiconductors; bandgap determines the photon energies at which transitions between bands can occur. Consequently, the tuning of the bandgap can enable the encoding of spectral information into a photoresponse vector. Miniaturized spectrometers have been demonstrated based on bandgap tuning by varying

Downloaded from https://www.science.org at University of Texas Dallas on May 11, 2023

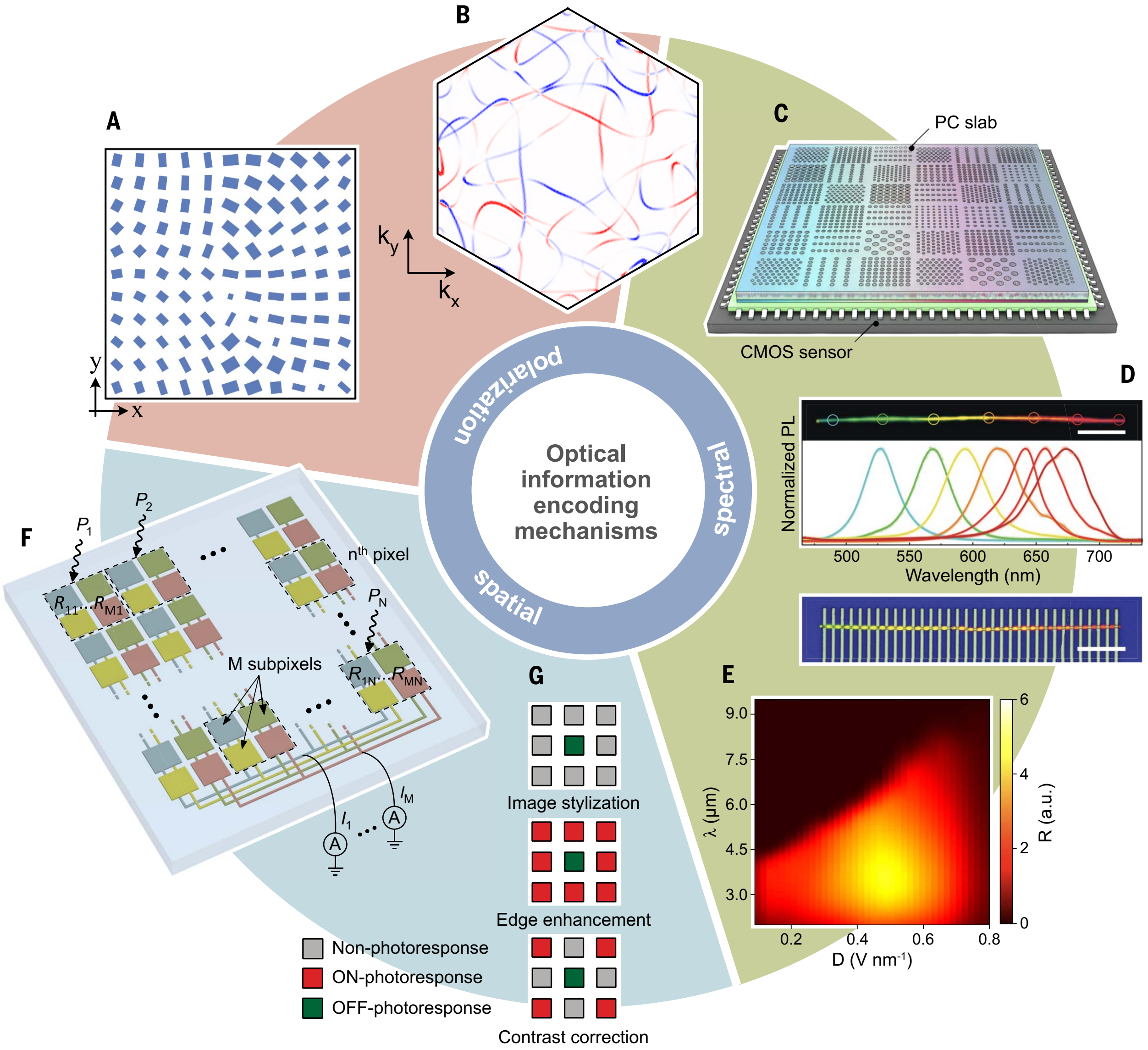


Fig. 2. Information encoding mechanisms. (**A**) Encoding the polarization information of light using a grating matrix with different physical geometries. [Adapted with permission from (*66*)] (**B**) Calculated distribution of the integrand for computing shift current of TDBG in the moiré Brillouin zone. Its tunability enables the encoding of the polarization, wavelength, and power information of mid-infrared light using the moiré quantum geometry of TDBG. [Adapted by permission from Springer Nature Customer Service Center GmbH, Springer *Nature* (*47*), copyright (2022)] (**C**) Encoding the spectral information of light using an array of photonic crystals (PCs) with different geometric features. [Adapted by permission from Springer Nature Customer Service Center GmbH, Springer *Nature* (*19*), copyright 2019)] (**D**) Bandgap tuning by varying the chemical composition for encoding spectral information. The photoluminescence (PL) spectra taken in different locations of a CdS_xSe_{1-x} nanowire with a graded composition are shown at the top. The shift in peak wavelength indicates the varying bandgap along the wire. A fabricated single-nanowire spectrometer is shown at the bottom. [Adapted with permission from (*20*)] (**E**) Spectral responsivity of a reconfigurable black phosphorus sensor under different biasing displacement fields. [Adapted by permission from Springer Nature Customer Service Center GmbH, Springer *Nature* (*25*), copyright (2021)] (**F**) Photodetector array for encoding optical images with *n* pixels into *m* electrical outputs. Each subpixel is reconfigurable by two split gates, which are biased with voltages of opposite polarities. [Adapted by permission from Springer Nature Customer Service Center GmbH, Springer *Nature* (*37*), copyright (2020)] (**G**) Photoresponsivity distributions of a reconfigurable pixel array for simultaneous image capture and processing (image stylization, edge enhancement, and contrast reduction). [Adapted with permission from (*39*)]

the chemical compositions of materials or the electrical displacement fields (*13*, *20*, *25*). In one case, different chemical compositions were introduced within a single cadmium sulfide selenide (CdS_xSe_{1-x}) nanowire to encode the spectral information of the incident light (*20*). The top panel of Fig. 2D shows the photoluminescence spectra taken at different locations along such a single nanowire, and the bottom panel illustrates a fabricated single-nanowire spectrometer. The spectral information of the light is encoded into the photoresponse vector that is measured along the wire. Moreover, the chemical composition and the dimension of quantum dots can be tuned together to cover a broad spectral range, as demonstrated in a quantum dot spectrometer (*13*).

An electric field can tune the absorption edge of a bulk semiconductor, which is well known as the Franz-Keldysh effect (*71*). In semiconductor quantum wells, such absorption

Downloaded from https://www.science.org at University of Texas Dallas on May 11, 2023

tuning is more pronounced and is referred to as the "quantum confinement Stark effect" (*72*–*75*). Recently, a single on-chip black phosphorus device has been shown to function as a mid-infrared spectrometer within the 2- to 9-μm wavelength range (*25*). The bandgap of ~10-nm-thick black phosphorus was tuned by an external electric field, by using the Stark effect for the encoding of spectral information. Figure 2E shows the spectral responsivity of a reconfigurable black phosphorus sensor (*25*) in which the cutoff wavelength has been extended to around 9 μm under a biasing displacement field of 0.8 V/nm. This demonstration has inspired further demonstrations of spectrometers based on single, reconfigurable photodetectors (*26*, *27*). In these works, the electric field does not appreciably tune the bandgap but rather changes the spectral response of the devices by adjusting the relative band alignment (*26*) or ion migration properties (*27*). Finally, metasurfaces on graphene exhibit strong tunability in the mid- and far-infrared wavelength regimes (*76*, *77*). Spectroscopy has also been demonstrated by combining such reconfigurable metasurfaces with discrete infrared photodetectors (*78*). In this case, the encoding process is realized by tuning the reflection spectra of the metasurfaces with an external electric field.

Encoding optical images with reconfigurable detector arrays

In addition to spectral and polarization information, the spatial variations of light intensity (optical images) can also be encoded in the photoresponse. To this end, an array of reconfigurable photodetectors as an image sensor and an artificial neural network have been used for ultrafast machine vision (*37*). The image-encoding process captures the spatial features directly and reduces the transmission bandwidth requirements. Figure 2F shows an illustration of the device, which consists of n photoactive pixels arranged in a 2D array, with each pixel divided into m subpixels (*37*). Each subpixel is composed of a WSe_2 photodiode whose responsivity can be reconfigured by two split gates. Enabled by the reconfigurability on the subpixel level, the sensor can be trained to encode optically projected images into an m-dimensional output, as will be further discussed in the section "Roles of neural networks in optical sensing."

The concept of capturing spatial features directly in the light-detection process has been used in several other works (*38*–*46*). For example, Fig. 2G illustrates three configurations of the responsivity matrix of a 3-pixel–by–3-pixel sensor array for image stylization, edge detection, and contrast correction in (*39*). Also in this case, different geometric features can be captured directly in the imaging process with different configurations of the device response, thus eliminating the need for subsequent computational image processing steps. In another example (*79*), a large fraction of the sensor elements in an imaging device were physically combined into several "superpixels" that extended over the entire surface area of the chip. For a given pattern recognition task, their optimal shapes were determined by using a machine-learning algorithm from training data. Classification of optically projected images on an ultrafast time scale and with an enhanced dynamic range was demonstrated.

General considerations for information encoding

We have shown that there are a number of pathways to realize information encoding. This naturally leads to the following question: What are the essential requirements for constructing a good encoder and a measurement channel? To reconstruct the physical information in $\mathcal{R}^n$ or to capture the features of interest directly from the m-dimensional photoresponse, degeneracy is not desirable. When multiple points in $\mathcal{R}^n$ are mapped to the same point in $\mathcal{R}^m$, loss-less reconstruction is no longer possible. Indeed, it is the degeneracy in their photoresponse that hinders conventional nonreconfigurable detectors from sensing richer information of unknown light because light beams with different combinations of physical properties (i.e., power, spectrum, polarization, angular momentum, geometric features, and so on) can yield the same output signal. The reconfigurability can eliminate degeneracy by increasing the dimensionality of the response. By configuring the geometric features, spectral response, and quantum geometric properties of the sensor, the tunability in its optical response can map different points in $\mathcal{R}^n$ to distinct points in $\mathcal{R}^m$. Noise introduced in the measurement process may, however, increase the possibility of overlapping originally distinctive points in $\mathcal{R}^m$, which leads to potential degeneracy. Therefore, a sufficiently large signal-to-noise ratio is important in the encoding process. This observation is analogous to the channel capacity in coding theory, where capacity increases as noise decreases. Other than noise reduction, introducing redundancy in the measurements ($m > n$) may mitigate the degeneracy problem (*61*).

Decoding pathways

Although the photoresponse vector $\mathbf{x}$ itself may contain valuable information about certain features of $\mathbf{w}$, a decoder is generally needed to decipher and reconstruct the original physical information to complete the sensing process. In this section, we discuss two general classes of models that are used in optical sensing to map the sensor response in $\mathcal{R}^m$ to the original physical information in $\mathcal{R}^n$ (or to capture some information of interest directly from $\mathbf{x}$): analytical (Fig. 3A) and data-driven (Fig. 3B). Analytical models require a comprehensive understanding of the encoder, whereas data-driven models usually use neural networks with experimental photoresponse data. In addition, other approaches exist that do not belong to either of the above two models but could represent alternative future pathways for decoding optical information. For example, randomly initialized neural networks without training have been shown to be effective in image generation and restoration (*80*, *81*), and an analytical algorithm followed by a convolutional neural network has been used to solve inverse problems (*82*).

For an exemplary illustration of an analytical decoding process, let us consider the following method, which is widely used in spectral sensing. In the linear response regime under optical excitation with spectrum P_λ, the response of a photodetector can be written as $I_S = \int R_\lambda P_\lambda \mathrm{d}\lambda$, where R_λ is the spectral responsivity (*13*–*15*, *18*, *20*, *21*). The light spectrum and the spectral responsivity can be represented in a vector space $\mathcal{R}^n$ by two vectors $\mathbf{p}_\lambda$ and $\mathbf{r}_\lambda$, respectively. The photoresponse I_S is then the inner product of these two vectors, $I_S = \mathbf{r}_\lambda^{\mathrm{T}} \mathbf{p}_\lambda$. [Similar equations hold for other linear optical properties, for example, the spatial intensity variation, $I_S = \mathbf{r}_x^{\mathrm{T}} \mathbf{p}_x$, where $\mathbf{p}_x$ represents a flattened optical image and $\mathbf{r}_x$ is a spatially varying photoresponsivity (*37*, *79*).]

Consider an optical sensor consisting of $m = n$ states, S_i, where $i = 1, 2, \ldots, n$. These n states, or measurements, may be realized by using n different subelements in the sensor or by n different operational modes of a single reconfigurable sensor. In either case, they can be represented as n inner products discussed above or as a matrix-vector product with photoresponsivity matrix $\mathbf{R}$, $\mathbf{i}_S = \mathbf{R}\mathbf{p}_\lambda$. Here, $\mathbf{R}$ is an n-by-n matrix, and its element at ith row and jth column, $R_{S_{i,j}}$, represents the discrete photoresponsivity at wavelength λ_j in state S_i; $\mathbf{i}_S$ denotes the discretized photoresponse vector $(I_{S_1}, I_{S_2}, \ldots I_{S_n})^T$; and the spectrum is denoted as $\mathbf{p}_\lambda = (P_{\lambda_1}, P_{\lambda_2}, \ldots P_{\lambda_n})^T$. If the response matrix $\mathbf{R}$ is known, then the spectrum $\mathbf{p}_\lambda$ can be reconstructed using the measured photoresponse $\mathbf{i}_S$. However, direct reconstruction by calculating the inverse $\mathbf{R}^{-1}$ may lead to unsatisfactory results because $\mathbf{R}$ may be ill-conditioned and $\mathbf{i}_S$ may exhibit measurement noise or even errors. In recent demonstrations, this problem has been mitigated using adaptive regression methods with Tikhonov (*83*) or LASSO (least absolute shrinkage and selection operator) regularizations (*84*) by minimizing a cost function, cost = $\|\mathbf{R}\mathbf{p}_\lambda - \mathbf{i}_S\|^2 + \alpha w(\mathbf{p}_\lambda)$. Here, α is a parameter that controls the regularization strength and $w(\mathbf{p}_\lambda)$ is a penalty function. These regularization approaches allow us to alleviate the negative effects of an ill-conditioned matrix $\mathbf{R}$ and measurement noise. In the general case, $n \neq m$, adaptive regression methods can still

Downloaded from https://www.science.org at University of Texas Dallas on May 11, 2023

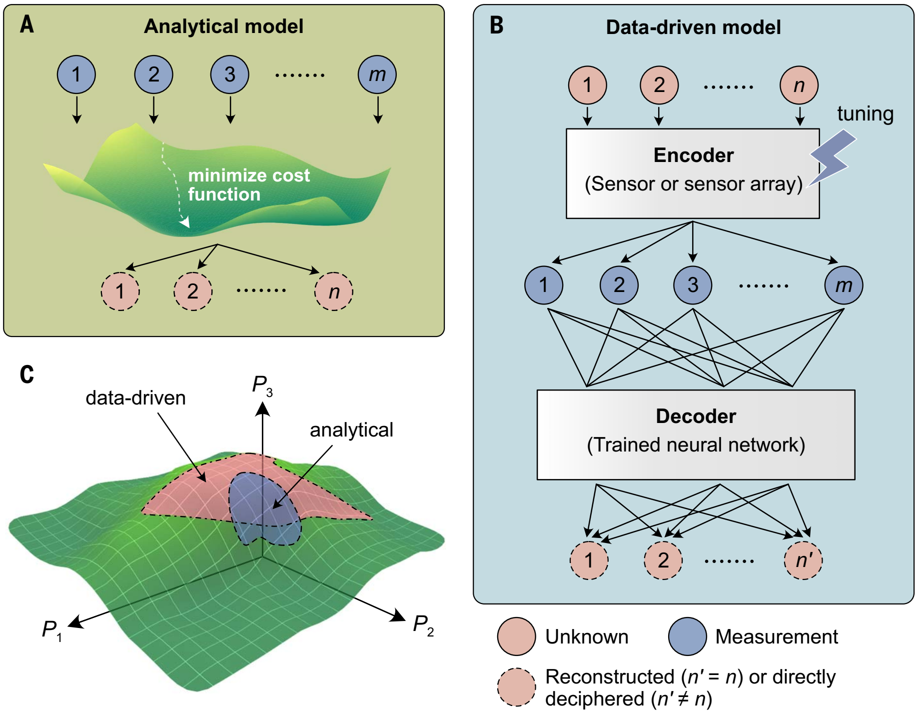


Fig. 3. Information decoding pathways. (**A**) Schematic of using an analytical approach to extract *n*-dimensional information from an *m*-dimensional photoresponse vector if the encoding process (not shown for simplicity) can be modeled explicitly. (**B**) Schematic of using a data-driven model in sensing. A reconfigurable device (or an array thereof) is used as an encoder to generate an *m*-dimensional photoresponse vector, and a trained neural network is used as a decoder to decipher the *n*-dimensional information ($n' = n$). The sensor itself can also act as (part of) a neural network. Moreover, desirable features of the *n*-dimensional information may be directly extracted from the *m*-dimensional photoresponse. (**C**) Schematic of a manifold in a conceptual, 3D parameter space. For analytical models, usually a smaller parameter space (blue surface) is required to capture the information, whereas a larger parameter space (red surface) is required for data-driven models.

be applied to compute a solution that minimizes the cost function.

Unlike analytical models, data-driven sensing generally consists of two steps. First, the model, usually a neural network, needs to be trained. This may occur on-chip or off-chip, in a supervised or an unsupervised or self-supervised manner (37–42, 47). In the former case, both a set of sensor inputs $\mathbf{w}$ and their photoresponse vectors $\mathbf{x}$ (or directly deciphered information, e.g., classification results) are provided. In the latter, an efficient representation (encoding) is learned from a set of inputs $\mathbf{w}$ alone, and the decoder attempts to reproduce at its output the original information, $\mathbf{w}' \approx \mathbf{w}$. After training, the neural network can then be leveraged to decipher the unknown information based on the measured photoresponse vector $\mathbf{x}$.

Data-driven models have several distinctive advantages that make them suitable for advanced sensing applications. First, data-driven models can be used as decoders to exploit existing experimental results, even when analytical models are not accessible. As illustrated in Fig. 3B, after the neural network is trained, the response can be interpreted without involving any specific mathematical relation. Second, different types of physical information can be deciphered simultaneously, as long as the training process takes them into account. Third, it is possible to realize functionality multiplexing because the outputs of the data-driven models are not limited to specific physical properties of light. For example, imaging and classification functions can be combined using trained neutral networks, which substantially reduces the complexity of the overall system.

At the same time, data-driven models require the acquisition of sufficient training data, which need to be correctly labeled in the case of supervised learning. If their characteristics do not change substantially during operation, then the sensors only need to be trained or calibrated once by the manufacturers without the end users having to go through this process. Data augmentation methods, such as interpolation and data synthesis, can be used to expand training datasets (85). In addition to the initial training of the model, recalibration during operation can also be applicable to sensors in both analytical and data-driven models. Choosing reliable references is crucial for deploying sensors in recalibration because laboratory-level calibration may not be available. Good references should have specific and stable features, and measuring them will provide enough information for recalibration. Examples include checkerboard and US Air Force (USAF) 1951 targets for imaging as well as elemental and molecular spectral lines for spectroscopy. Recalibration should focus on parameters directly related to drifting and degradation, which require a comprehensive understanding of the sensor's physical properties.

Manufacturing and environmental variations and measurement noise are further issues that need to be considered. Sensors may be sensitive to manufacturing variability and environmental conditions, such as temperature, humidity, and stray light. Advanced packaging schemes can increase tolerance to these conditions. The effects of these variations can be compensated numerically, if well understood. Measurement noise can be minimized by optimizing encoding processes and taking into account the sensor's physical properties and sensing requirements. Designing application-specific encoding strategies is critical to achieve both efficiency and accuracy. For example, focusing on operational states that are strongly affected by targeted spectral features can improve the performance of spectral sensing.

Regardless of the model used, the dimensionality m of the photoresponse vector is determined by the measurement parameter space. The parameters used to characterize the photoresponse vector can be diverse. For example, in tunable dual-gate sensors (25, 47), top (V_{TG}) and bottom (V_{BG}) gate biases are typical parameters that together form a 2D parameter space. Bias voltage in the photocurrent generation path, ambient temperature, load applied to the sensor, and external magnetic field may also be among the parameters for photocurrent measurements, depending on the sensing application. A conceptual, 3D parameter space is shown in Fig. 3C, as illustrated by axes P_1, P_2, and P_3. In practice, the parameter space can be reduced by a set of constraints to a manifold (green surface in Fig. 3C) on which the photoresponse is measured. When analytical models are applied, the understanding on the sensor is usually extensive. As a result, a smaller parameter space can be used. For example, in a previously demonstrated dual-gate black phosphorus spectrometer, it is known that at charge-neutrality, the photoresponse is highest and the bandgap tuning is effective (25). As a result, it is not necessary to perform the photoresponse measurements across the entire 2D parameter space of V_{TG} and V_{BG}, but they can instead be performed in a 1D line along which V_{TG} and V_{BG} collectively induce no net doping. In a data-driven model, the understanding of the sensor does not necessarily need to be as comprehensive. In this case, usually a larger parameter space is needed

Downloaded from https://www.science.org at University of Texas Dallas on May 11, 2023

to fully capture the information of unknown light, as illustrated in (*47*), in which a 2D parameter space of V_{TG} and V_{BG} was used. The larger (red) and smaller (blue) areas in Fig. 3C schematically represent these two types of parameter spaces, respectively.

Roles of neural networks in optical sensing

We now discuss potential roles of artificial neural networks in the information encoding and decoding in optical sensors. Autoencoders are a specific class of neural networks that are particularly promising for the sensing scheme introduced here. Figure 4A shows a schematic of the network (*37*). An autoencoder consists of two parts: an encoder that compresses the input data in a bottleneck layer with $m < n$ dimensions and a decoder that attempts to reproduce the original data at its output. Figure 2F shows an illustration of a device realization for image encoding (*37*). We emphasize, however, that the concept is not limited to imaging alone; it can also be applied to other optical sensing tasks, such as spectral measurements. The device consists of $n = 9$ photoactive pixels arranged in a 2D array, and each pixel is divided into $m = 3$ subpixels. Each subpixel consists of a WSe_2 photodiode whose photoresponsivity can be configured to adjust the synaptic weights. By interconnecting the subpixels, an integrated neural network can be formed in which the encoder is the optical sensor itself and the decoder is the external computer programs. Figure 4B shows the operation of a device after a training process that is based on backpropagation (*37*). The encoder translates the projected images (letters “n,” “v,” and “z”) into an output current vector, which is converted by a nonlinearity into an activation code and finally reconstructed into the original image by the decoder. It is noticeable in Fig. 4B that the activation codes are binary. This is a consequence of the training process, in which Gaussian noise was injected to learn binary representations. Such a device may be operated as a binary-hashing autoencoder, which eliminates the need for analog-to-digital conversion before signal reconstruction. To implement a deep autoencoder, additional hidden layers can be added to both the encoder and the decoder to deepen the network and increase its complexity and performance. Although this is straightforward on the (software) decoder side, on the (hardware) encoder side, it is more elaborate but could be achieved, for example, by converting the output currents to voltages that are then fed into a memristor crossbar (*86*, *87*). However, as demonstrated in Fig. 4C, comparable performance can be achieved by keeping a single layer for the encoder and only increasing the number of layers of the decoder.

Other neural network architectures can also be implemented using a similar device structure. A machine vision processor was developed to operate as a convolutional neural network by integrating 1024 MoS_2 photo-field effect transistors in a crossbar structure, and a classification of digits from the Modified National Institute of Standards and Technology (MNIST) dataset was demonstrated (*38*). Figure 4D illustrates the working principles of another classifier presented in (*39*) by implementing a convolutional neural network with a prototypical 3-pixel–by–3-pixel sensor. Here, the photoresponsivity of each pixel is reconfigurable, and the total photocurrent represents the convolution of the image and the responsivity matrix. Binary figures representing the letters “n,” “j,” and “u” were used during the training to obtain the responsivity matrix of each letter. A testing classification accuracy of 100% was achieved by using the weighted average of the convolutional kernel. In a third example (*42*), an array of black phosphorus programmable phototransistors that can be programmed with 5-bit precision was used to implement an in-sensor convolutional neural network.

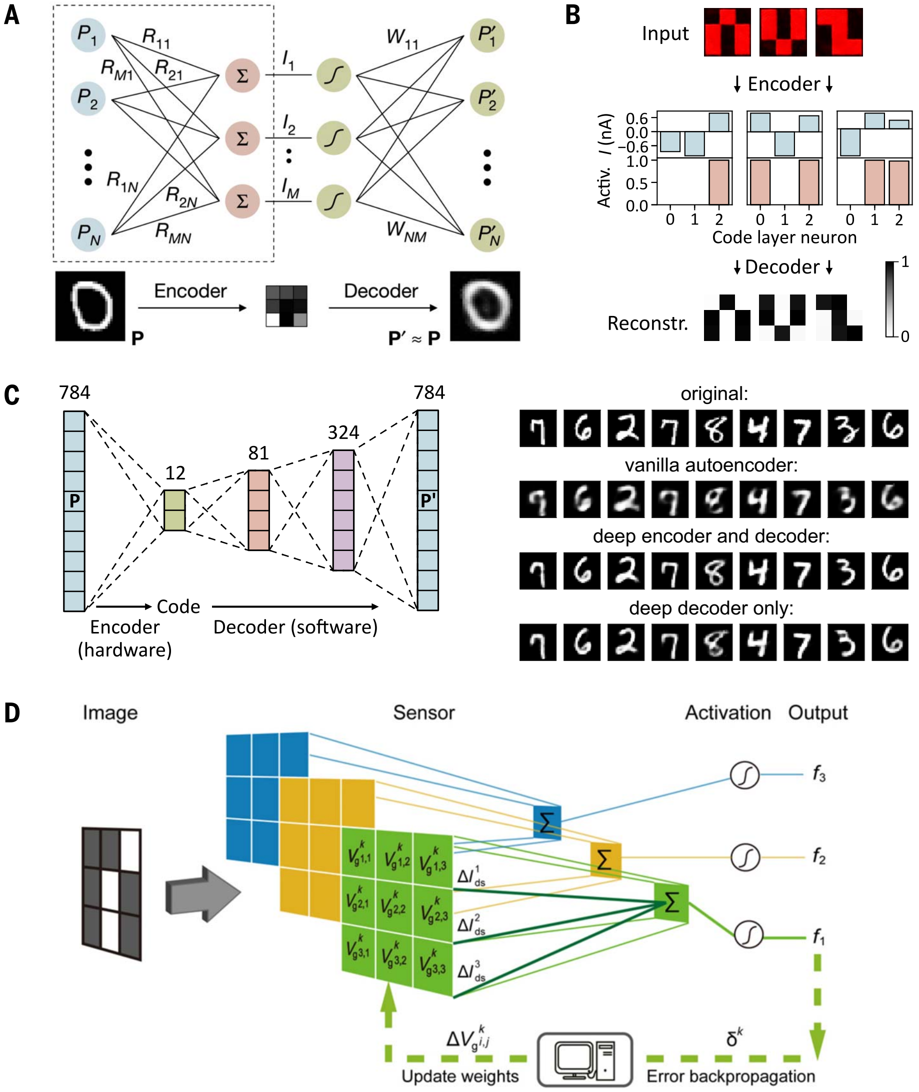


Fig. 4. Machine vision using a reconfigurable sensor array. (**A**) Illustration of an autoencoder with a single hidden layer. The bottom shows the encoding and decoding of a letter from the MNIST database. (**B**) Operation of an autoencoder based on a reconfigurable 9-pixel WSe_2 sensor array. The sensor array acts as an encoder that translates images into current codes that can later be reconstructed into the original image by an external decoder. [(A) and (B) are adapted by permission from Springer Nature Customer Service Center GmbH, Springer *Nature* (*37*), copyright (2020)] (**C**) Illustration of an autoencoder with a deep decoder (left) and MNIST image ($n = 784$) reconstruction ($m = 12$) using different types of autoencoders (right). (**D**) Convolutional neural network realized by a reconfigurable 9-pixel retinomorphic vision sensor. [Adapted with permission from (*39*)]

In addition to the image recognition and processing functions described above, it is also possible to detect multiple physical properties

Downloaded from https://www.science.org at University of Texas Dallas on May 11, 2023

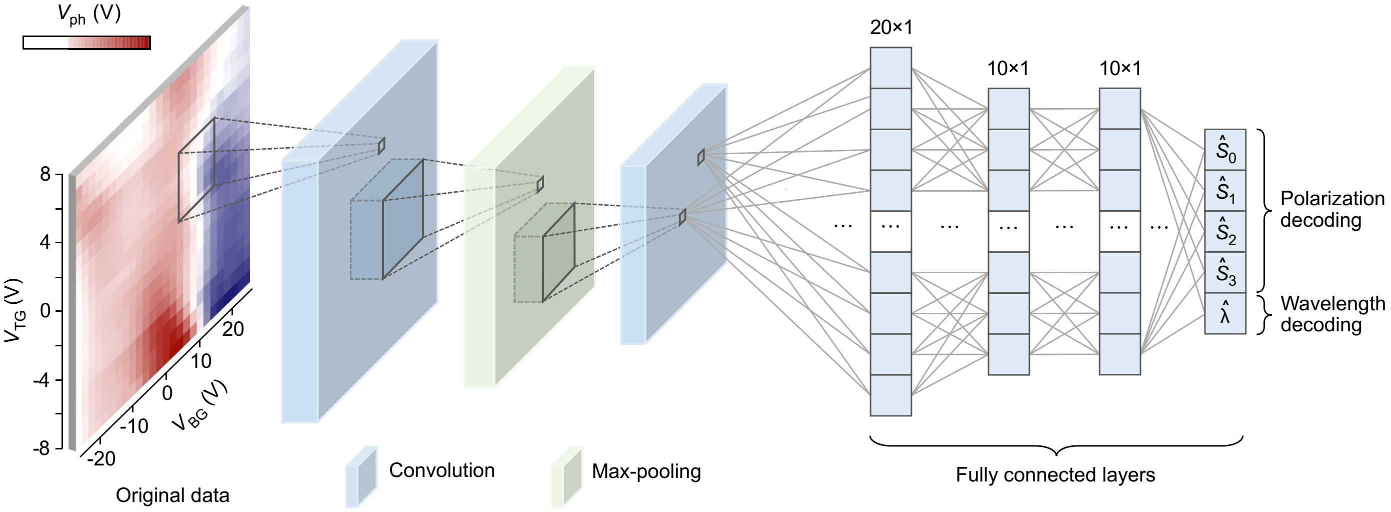


Fig. 5. Deep neural network polarimetry and wavelength detection. Schematic of the convolutional neural network used for the demonstration of the polarimeter and wavelength detection. The input layer is the measured 20-pixel–by–26-pixel photovoltage (V_{ph}) mapping (leftmost panel), and the output is a five-element vector, $\left(\hat{S}_0, \hat{S}_1, \hat{S}_2, \hat{S}_3, \hat{\lambda}\right)$. The mapping, in which the polarization and wavelength information of the incident 5-μm light is encoded, consists of 20 pixels by 26 pixels, which corresponds to 20 values of top-gate voltage (V_{TG}) and 26 values of back-gate voltage (V_{BG}). The hidden layers consist of the first convolution layer, the max-pooling layer, the second convolution layer, and three subsequent fully connected layers. Stokes parameters can be directly calculated using $\left(\hat{S}_0, \hat{S}_1, \hat{S}_2, \hat{S}_3\right)$. A wavelength label $\hat{\lambda} = \mp 1$ corresponds to 5 and 7.7 μm, respectively. [Adapted by permission from Springer Nature Customer Service Center GmbH, Springer *Nature* (47), copyright (2022)]

simultaneously using a reconfigurable sensor. As discussed in the section "Information-encoding mechanisms," by using the BPVE, the polarization state, wavelength, and power information can be encoded into a 2D photoresponse map (47), as shown in the leftmost panel of Fig. 5. Although the mechanisms of the BPVE are well understood within the framework of quantum geometry, precise analytical modeling of the measured photoresponse is not feasible because of the extrinsic complexities of the moiré system, such as finite temperature, unintentional strain, and twist-angle disorder (67). Instead, a convolutional neural network can be trained as the decoder by using a large number of such 2D mappings from excitation lights with known physical properties. The trained convolutional neural network can then be used to decipher the 2D mapping to reveal the wavelength, power, and polarization state of an unknown light (47), as illustrated in Fig. 5. We expect deep neural networks to play an increasingly important role as the sensing tasks become more complex and demanding.

Discussion and outlook

Emerging opportunities in deep optical sensing

As noted above, a single TDBG sensor can simultaneously detect the wavelength, intensity, and polarization state of unknown light in the mid-infrared regime. Extending this capability to other wavelength ranges and enabling the detection of other physical properties, such as angular momentum (88–91), will further empower this sensing scheme. Innovative materials will be needed to demonstrate deep optical sensing beyond the mid-infrared. Moreover, the tunable BPVE, which is central to the TDBG reconfigurable sensor concept, is a second-order optical effect (47). The regular photovoltaic and photoconductive responses can also be reconfigurable (37–42). The use of reconfigurable linear and higher-order photoresponses together may further enhance the capabilities and improve the performance of deep optical sensing.

Another future direction is to expand the capability of functionality multiplexing. It has been shown that an array of reconfigurable photodetectors can realize image-encoding and -classification functions based on artificial neural networks (37). However, the device can only handle simple images because it is limited by the low resolution of the array and the low complexity of the neural network architecture. The construction of reconfigurable sensor arrays with higher resolution and more layers will enable the use of deep neural networks in conjunction with enhanced imaging capabilities for more challenging machine-vision tasks. Simultaneous encoding of both spectral and spatial information may lead to a new generation of high-throughput hyperspectral imaging systems.

Identifying innovative sensing materials and mechanisms

As sensors become ubiquitous, it is highly desirable to continuously reduce their size to enable on-chip integration. To achieve this goal, reconfigurability is key, as emphasized throughout this Review. Miniaturized sensors that can perform a wide range of different tasks have been demonstrated mostly using 2D materials such as black phosphorus (25), transition metal dichalcogenides (37), moiré graphene (47), and perovskites (27). Research on other material systems to realize reconfigurability will likely extend the operational spectral range and enable new functionalities. Van der Waals heterostructures, for example, represent a diverse spectrum of material systems with strong tunability that can interact with light in the wavelength range from microwave to ultraviolet. External electric fields can tune not only the doping (or Fermi energy) and the bandgap of the constituent materials in the heterostructure but also the relative band alignments between different layers (26) and the quantum geometric properties (47, 57). As a result, optical transitions within and between the layers, as well as nonlinear optical effects, can all be reconfigured by electric fields, which provides ample opportunities for the realization of deep optical sensing in a broad wavelength range. Moreover, conventional thin-film semiconductors such as silicon-germanium and III-V quantum wells also exhibit tunability under electric fields (72, 73, 92), which makes it feasible to build reconfigurable sensors based on highly mature semiconductor platforms. A recently demonstrated silicon reconfigurable imager has shown the potential of using silicon simultaneously for imaging and in-sensor data processing in the visible spectral range (93). Compressive sensing and imaging can also benefit from device reconfigurability (46, 65).

Downloaded from https://www.science.org at University of Texas Dallas on May 11, 2023

Integration of sensing and computing functionalities

Reconfigurable sensors may generate more data than conventional sensors because of their multiple operational states. Reading and processing these large amounts of data can be challenging. Developing information processing schemes with integrated memory cells near or within the sensors may thus be crucial. For example, a reconfigurable integrated sensor array, based on a van der Waals heterostructure, has been developed that incorporates sensing, memory, and computing functions (*41*). This sensor array exhibits nonvolatile negative and positive photoresponses, which are used for motion detection, and external neural networks can be combined with the sensor to enable more advanced functionalities. A recent article (*94*) has extensively discussed the possible strategies for performing in- or near-sensor computing. Deep optical sensing is expected to benefit tremendously from this related field. Further, it is expected that the hybrid integration of reconfigurable sensors made from 2D materials, perovskites, and other thin-film semiconductors with silicon electronics for data processing may pave the way for a new generation of deep sensing technologies.

Optical computing is another promising pathway for high-throughput information processing, which benefits both encoders and decoders. Optical computing can enable higher degrees of freedom in encoder designs by directly processing the optical signals and high-throughput decoding through deep neural networks in hardware format, owing to the high bandwidth of optical modulation and high speed of light. Functions such as computer vision and language processing have been demonstrated based on optical deep neural networks (*95*, *96*).

Leveraging the latest developments in machine learning

The development of deep optical sensing schemes with reconfigurable sensors provides a singular opportunity to test and exploit the latest developments in machine learning. For example, a generative-adversarial network (GAN) has been used to enable compressed sensing without assuming sparsity (*97*). A GAN typically consists of two competing neural networks: a generator and a discriminator (*49*). The generator is trained to produce vectors from random noise to fool the discriminator, whereas the discriminator attempts to distinguish the generated vectors from existing datasets. By properly training these two neural networks, the generator learns the distribution of existing data and eventually generates vectors that possess the features of the dataset. By incorporating prior experience or knowledge, the reconstruction error can be made small even with a limited number of measurements. We expect that GANs can be applied to improve the deep sensing performance in terms of the reduction of required measurements. Another example is the long short-term memory (LSTM), which can be used together with reconfigurable optical sensors to detect ultrafast events in computer vision and chemical reactions (*49*).

Establishing mathematical guidelines for encoding

One challenge in information encoding is to minimize the required number of measurement states m and to identify an optimal encoding strategy. If fewer measurements can be performed without compromising sensing performance, then not only can the acquisition speed be increased, but the data processing requirements in subsequent steps can also be reduced. By contrast, by choosing a large m, redundancy can be introduced, which reduces the probability of degeneracy of the photoresponse. Therefore, a set of mathematical guidelines are needed to bridge the gap between these two conflicting requirements.

In conclusion, optical sensing will benefit tremendously from the latest developments in device technology, materials science, condensed matter physics, and machine learning. Future sensors are likely to be highly compact, reconfigurable, multifunctional, and intelligent, and they will find applications in medical imaging, environmental monitoring, infrared astronomy, and many other areas of our daily lives, especially in the mobile domain.

REFERENCES AND NOTES

1. D. C. D. Pocock, Sight and knowledge. *Trans. Inst. Br. Geogr.* **6**, 385–393 (1981). doi: 10.2307/621875
2. Z. Yang, T. Albrow-Owen, W. Cai, T. Hasan, Miniaturization of optical spectrometers. *Science* **371**, eabe0722 (2021). doi: 10.1126/science.abe0722; pmid: 33509998
3. L. Gao, Y. Qu, L. Wang, Z. Yu, Computational spectrometers enabled by nanophotonics and deep learning. *Nanophotonics* **11**, 2507–2529 (2022). doi: 10.1515/nanoph-2021-0636
4. T. Yang *et al.*, Miniature spectrometer based on diffraction in a dispersive hole array. *Opt. Lett.* **40**, 3217–3220 (2015). doi: 10.1364/OL.40.003217; pmid: 26125406
5. B. Redding, S. Fatt Liew, Y. Bromberg, R. Sarma, H. Cao, Evanescently coupled multimode spiral spectrometer. *Optica* **3**, 956–962 (2016). doi: 10.1364/OPTICA.3.000956
6. P. Edwards *et al.*, Smartphone based optical spectrometer for diffusive reflectance spectroscopic measurement of hemoglobin. *Sci. Rep.* **7**, 12224 (2017). doi: 10.1038/s41598-017-12482-5; pmid: 28939898
7. M. Faraji-Dana *et al.*, Compact folded metasurface spectrometer. *Nat. Commun.* **9**, 4196 (2018). doi: 10.1038/s41467-018-06495-5; pmid: 30305616
8. W. Hartmann *et al.*, Waveguide-integrated broadband spectrometer based on tailored disorder. *Adv. Opt. Mater.* **8**, 1901602 (2020). doi: 10.1002/adom.201901602
9. A. V. Velasco *et al.*, High-resolution Fourier-transform spectrometer chip with microphotonic silicon spiral waveguides. *Opt. Lett.* **38**, 706–708 (2013). doi: 10.1364/OL.38.000706; pmid: 23455272
10. D. M. Kita *et al.*, High-performance and scalable on-chip digital Fourier transform spectroscopy. *Nat. Commun.* **9**, 4405 (2018). doi: 10.1038/s41467-018-06773-2; pmid: 30353014
11. S. N. Zheng *et al.*, Microring resonator-assisted Fourier transform spectrometer with enhanced resolution and large bandwidth in single chip solution. *Nat. Commun.* **10**, 2349 (2019). doi: 10.1038/s41467-019-10282-1; pmid: 31138800
12. D. Pohl *et al.*, An integrated broadband spectrometer on thin-film lithium niobate. *Nat. Photonics* **14**, 24–29 (2020). doi: 10.1038/s41566-019-0529-9
13. J. Bao, M. G. Bawendi, A colloidal quantum dot spectrometer. *Nature* **523**, 67–70 (2015). doi: 10.1038/nature14576; pmid: 26135449
14. E. Huang, Q. Ma, Z. Liu, Etalon array reconstructive spectrometry. *Sci. Rep.* **7**, 40693 (2017). pmid: 28074883
15. B. Craig, V. R. Shrestha, J. Meng, J. J. Cadusch, K. B. Crozier, Experimental demonstration of infrared spectral reconstruction using plasmonic metasurfaces. *Opt. Lett.* **43**, 4481–4484 (2018). doi: 10.1364/OL.43.004481; pmid: 30211895
16. A. Tittl *et al.*, Imaging-based molecular barcoding with pixelated dielectric metasurfaces. *Science* **360**, 1105–1109 (2018). doi: 10.1126/science.aas9768; pmid: 29880685
17. Y. Zhu, X. Lei, K. X. Wang, Z. Yu, Compact CMOS spectral sensor for the visible spectrum. *Photon. Res.* **7**, 961–966 (2019). doi: 10.1364/PRJ.7.000961
18. J. Meng, J. J. Cadusch, K. B. Crozier, Detector-only spectrometer based on structurally colored silicon nanowires and a reconstruction algorithm. *Nano Lett.* **20**, 320–328 (2020). doi: 10.1021/acs.nanolett.9b03862; pmid: 31829611
19. Z. Wang *et al.*, Single-shot on-chip spectral sensors based on photonic crystal slabs. *Nat. Commun.* **10**, 1020 (2019). doi: 10.1038/s41467-019-08994-5; pmid: 30833569
20. Z. Yang *et al.*, Single-nanowire spectrometers. *Science* **365**, 1017–1020 (2019). doi: 10.1126/science.aax8814; pmid: 31488686
21. X. Zhu *et al.*, Broadband perovskite quantum dot spectrometer beyond human visual resolution. *Light Sci. Appl.* **9**, 73 (2020). doi: 10.1038/s41377-020-0301-4; pmid: 32377335
22. J. Zhang, X. Zhu, J. Bao, Denoising autoencoder aided spectrum reconstruction for colloidal quantum dot spectrometers. *IEEE Sens. J.* **21**, 6450–6458 (2020). doi: 10.1109/JSEN.2020.3039973
23. C. Brown *et al.*, Neural network-based on-chip spectroscopy using a scalable plasmonic encoder. *ACS Nano* **15**, 6305–6315 (2021). doi: 10.1021/acsnano.1c00079; pmid: 33543919
24. K. D. Hakkel *et al.*, Integrated near-infrared spectral sensing. *Nat. Commun.* **13**, 103 (2022). doi: 10.1038/s41467-021-27662-1; pmid: 35013200
25. S. Yuan, D. Naveh, K. Watanabe, T. Taniguchi, F. Xia, A wavelength-scale black phosphorus spectrometer. *Nat. Photonics* **15**, 601–607 (2021). doi: 10.1038/s41566-021-00787-x
26. W. Deng *et al.*, Electrically tunable two-dimensional heterojunctions for miniaturized near-infrared spectrometers. *Nat. Commun.* **13**, 4627 (2022). doi: 10.1038/s41467-022-32306-z; pmid: 35941126
27. L. Guo *et al.*, A single-dot perovskite spectrometer. *Adv. Mater.* **34**, e2200221 (2022). doi: 10.1002/adma.202200221; pmid: 35706366
28. H. H. Yoon *et al.*, Miniaturized spectrometers with a tunable van der Waals junction. *Science* **378**, 296–299 (2022). doi: 10.1126/science.add8544; pmid: 36264793
29. L. Tong *et al.*, Stable mid-infrared polarization imaging based on quasi-2D tellurium at room temperature. *Nat. Commun.* **11**, 2308 (2020). doi: 10.1038/s41467-020-16125-8; pmid: 32385242
30. J. Wei, C. Xu, B. Dong, C.-W. Qiu, C. Lee, Mid-infrared semimetal polarization detectors with configurable polarity transition. *Nat. Photonics* **15**, 614–621 (2021). doi: 10.1038/s41566-021-00819-6
31. J. Wei *et al.*, Geometric filterless photodetectors for mid-infrared spin light. *Nat. Photonics* **17**, 171–178 (2022). doi: 10.1038/s41566-022-01115-7
32. F. Yesilkoy *et al.*, Ultrasensitive hyperspectral imaging and biodetection enabled by dielectric metasurfaces. *Nat. Photonics* **13**, 390–396 (2019). doi: 10.1038/s41566-019-0394-6
33. K. Monakhova, K. Yanny, N. Aggarwal, L. Waller, Spectral diffusercam: Lensless snapshot hyperspectral imaging with a spectral filter array. *Optica* **7**, 1298–1307 (2020). doi: 10.1364/OPTICA.397214
34. W. Zhang *et al.*, Deeply learned broadband encoding stochastic hyperspectral imaging. *Light Sci. Appl.* **10**, 108 (2021). doi: 10.1038/s41377-021-00545-2; pmid: 34035213
35. Z. Ballard, C. Brown, A. M. Madni, A. Ozcan, Machine learning and computation-enabled intelligent sensor design. *Nat. Mach. Intell.* **3**, 556–565 (2021). doi: 10.1038/s42256-021-00360-9
36. Y. Luo *et al.*, Design of task-specific optical systems using broadband diffractive neural networks. *Light Sci. Appl.* **8**, 112 (2019). doi: 10.1038/s41377-019-0223-1; pmid: 31814969

Downloaded from https://www.science.org at University of Texas Dallas on May 11, 2023

37. L. Mennel *et al*., Ultrafast machine vision with 2D material neural network image sensors. *Nature* **579**, 62–66 (2020). doi: 10.1038/s41586-020-2038-x; pmid: 32132692
38. H. Jang *et al*., An atomically thin optoelectronic machine vision processor. *Adv. Mater.* **32**, e2002431 (2020). doi: 10.1002/adma.202002431; pmid: 32700395
39. C.-Y. Wang *et al*., Gate-tunable van der Waals heterostructure for reconfigurable neural network vision sensor. *Sci. Adv.* **6**, eaba6173 (2020). doi: 10.1126/sciadv.aba6173; pmid: 32637614
40. T. Ahmed *et al*., Fully light-controlled memory and neuromorphic computation in layered black phosphorus. *Adv. Mater.* **33**, e2004207 (2021). doi: 10.1002/adma.202004207; pmid: 33205523
41. Z. Zhang *et al*., All-in-one two-dimensional retinomorphic hardware device for motion detection and recognition. *Nat. Nanotechnol.* **17**, 27–32 (2022). doi: 10.1038/s41565-021-01003-1; pmid: 34750561
42. S. Lee, R. Peng, C. Wu, M. Li, Programmable black phosphorus image sensor for broadband optoelectronic edge computing. *Nat. Commun.* **13**, 1485 (2022). doi: 10.1038/s41467-022-29171-1; pmid: 35304489
43. S. Wang *et al*., Networking retinomorphic sensor with memristive crossbar for brain-inspired visual perception. *Natl. Sci. Rev.* **8**, nwaa172 (2020). doi: 10.1093/nsr/nwaa172; pmid: 34691573
44. Q.-B. Zhu *et al*., A flexible ultrasensitive optoelectronic sensor array for neuromorphic vision systems. *Nat. Commun.* **12**, 1798 (2021). doi: 10.1038/s41467-021-22047-w; pmid: 33741964
45. C. Choi *et al*., Curved neuromorphic image sensor array using a MoS_2-organic heterostructure inspired by the human visual recognition system. *Nat. Commun.* **11**, 5934 (2020). doi: 10.1038/s41467-020-19806-6
46. L. Mennel, D. K. Polyushkin, D. Kwak, T. Mueller, Sparse pixel image sensor. *Sci. Rep.* **12**, 5650 (2022). doi: 10.1038/s41598-022-09594-y; pmid: 35383216
47. C. Ma *et al*., Intelligent infrared sensing enabled by tunable moiré quantum geometry. *Nature* **604**, 266–272 (2022). doi: 10.1038/s41586-022-04548-w; pmid: 35418636
48. M. M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, P. Vandergheynst, Geometric deep learning: Going beyond euclidean data. *IEEE Signal Process. Mag.* **34**, 18–42 (2017). doi: 10.1109/MSP.2017.2693418
49. I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning* (MIT Press, 2016).
50. Q. Ma, A. G. Grushin, K. S. Burch, Topology and geometry under the nonlinear electromagnetic spotlight. *Nat. Mater.* **20**, 1601–1614 (2021). doi: 10.1038/s41563-021-00992-7; pmid: 34127824
51. A. M. Cook, B. M Fregoso, F. de Juan, S. Coh, J. E. Moore, Design principles for shift current photovoltaics. *Nat. Commun.* **8**, 14176 (2017). doi: 10.1038/ncomms14176; pmid: 28120823
52. T. Morimoto, N. Nagaosa, Topological nature of nonlinear optical effects in solids. *Sci. Adv.* **2**, e1501524 (2016). doi: 10.1126/sciadv.1501524; pmid: 27386523
53. J. Sipe, A. Shkrebtii, Second-order optical response in semiconductors. *Phys. Rev. B* **61**, 5337–5352 (2000). doi: 10.1103/PhysRevB.61.5337
54. J. Ahn, G.-Y. Guo, N. Nagaosa, A. Vishwanath, Riemannian geometry of resonant optical responses. *Nat. Phys.* **18**, 290–295 (2022). doi: 10.1038/s41567-021-01465-z
55. E. Cohen *et al*., Geometric phase from Aharonov–Bohm to Pancharatnam–Berry and beyond. *Nat. Rev. Phys.* **1**, 437–449 (2019). doi: 10.1038/s42254-019-0071-1
56. Q. Ma *et al*., Direct optical detection of Weyl fermion chirality in a topological semimetal. *Nat. Phys.* **13**, 842–847 (2017). doi: 10.1038/nphys4146
57. S.-Y. Xu *et al*., Electrically switchable Berry curvature dipole in the monolayer topological insulator WTe_2. *Nat. Phys.* **14**, 900–906 (2018). doi: 10.1038/s41567-018-0189-6
58. G. B. Osterhoudt *et al*., Colossal mid-infrared bulk photovoltaic effect in a type-I Weyl semimetal. *Nat. Mater.* **18**, 471–475 (2019). doi: 10.1038/s41563-019-0297-4; pmid: 30833781
59. J. Ma *et al*., Nonlinear photoresponse of type-II Weyl semimetals. *Nat. Mater.* **18**, 476–481 (2019). doi: 10.1038/s41563-019-0296-5; pmid: 30833780
60. T. Akamatsu *et al*., A van der Waals interface that creates in-plane polarization and a spontaneous photovoltaic effect. *Science* **372**, 68–72 (2021). doi: 10.1126/science.aaz9146; pmid: 33795452
61. T. M. Cover, *Elements of Information Theory* (Wiley, 1999).
62. Y. August, A. Stern, Compressive sensing spectrometry based on liquid crystal devices. *Opt. Lett.* **38**, 4996–4999 (2013). doi: 10.1364/OL.38.004996; pmid: 24281493
63. D. L. Donoho, Compressed sensing. *IEEE Trans. Inf. Theory* **52**, 1289–1306 (2006). doi: 10.1109/TIT.2006.871582
64. Y. C. Eldar, G. Kutyniok, *Compressed Sensing: Theory and Applications* (Cambridge Univ. Press, 2012).
65. M. F. Duarte *et al*., Single-pixel imaging via compressive sampling. *IEEE Signal Process. Mag.* **25**, 83–91 (2008). doi: 10.1109/MSP.2007.914730
66. N. A. Rubin *et al*., Matrix Fourier optics enables a compact full-Stokes polarization camera. *Science* **365**, eaax1839 (2019). doi: 10.1126/science.aax1839; pmid: 31273096
67. C. N. Lau, M. W. Bockrath, K. F. Mak, F. Zhang, Reproducibility in the fabrication and physics of moiré materials. *Nature* **602**, 41–50 (2022). doi: 10.1038/s41586-021-04173-z; pmid: 35110759
68. J. Xiong, X. Cai, K. Cui, Y. Huang, J. Yang, H. Zhu, Z. Zheng, S. Xu, Y. He, F. Liu, X. Feng, W. Zhang, One-shot ultraspectral imaging with reconfigurable metasurfaces. arXiv:2005.02689 [physics.optics] (2020).
69. Z. Wang, Z. Yu, Spectral analysis based on compressive sensing in nanophotonic structures. *Opt. Express* **22**, 25608–25614 (2014). doi: 10.1364/OE.22.025608; pmid: 25401594
70. J. J. Cadusch, J. Meng, B. Craig, K. B. Crozier, Silicon microspectrometer chip based on nanostructured fishnet photodetectors with tailored responsivities and machine learning. *Optica* **6**, 1171–1177 (2019). doi: 10.1364/OPTICA.6.001171
71. S. L. Chuang, *Physics of Photonic Devices* (Wiley, 2012).
72. D. A. Miller *et al*., Band-edge electroabsorption in quantum well structures: The quantum-confined stark effect. *Phys. Rev. Lett.* **53**, 2173–2176 (1984). doi: 10.1103/PhysRevLett.53.2173
73. Y.-H. Kuo *et al*., Strong quantum-confined Stark effect in germanium quantum-well structures on silicon. *Nature* **437**, 1334–1336 (2005). doi: 10.1038/nature04204; pmid: 16251959
74. S. A. Empedocles, M. G. Bawendi, Quantum-confined stark effect in single CdSe nanocrystallite quantum dots. *Science* **278**, 2114–2117 (1997). doi: 10.1126/science.278.5346.2114; pmid: 9405345
75. C. Lin, R. Grassi, T. Low, A. S. Helmy, Multilayer black phosphorus as a versatile mid-infrared electro-optic material. *Nano Lett.* **16**, 1683–1689 (2016). doi: 10.1021/acs.nanolett.5b04594; pmid: 26901350
76. Y. Yao *et al*., Electrically tunable metasurface perfect absorbers for ultrathin mid-infrared optical modulators. *Nano Lett.* **14**, 6526–6532 (2014). doi: 10.1021/nl503104n; pmid: 25310847
77. Y. Yao *et al*., Broad electrical tuning of graphene-loaded plasmonic antennas. *Nano Lett.* **13**, 1257–1264 (2013). doi: 10.1021/nl3047943; pmid: 23441688
78. V. R. Shrestha *et al*., Mid- to long-wave infrared computational spectroscopy with a graphene metasurface modulator. *Sci. Rep.* **10**, 5377 (2020). doi: 10.1038/s41598-020-61998-w; pmid: 32214114
79. L. Mennel *et al*., A photosensor employing data-driven binning for ultrafast image recognition. *Sci. Rep.* **12**, 14441 (2022). doi: 10.1038/s41598-022-18821-5; pmid: 36002539
80. D. Ulyanov, A. Vedaldi, V. Lempitsky, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2018), pp. 9446–9454.
81. G. Mataev, P. Milanfar, M. Elad, DeepRED: Deep image prior powered by RED. arxiv:1903.10176 [cx.CV] (2019).
82. K. H. Jin, M. T. McCann, E. Froustey, M. Unser, Deep convolutional neural network for inverse problems in imaging. *IEEE Trans. Image Process.* **26**, 4509–4522 (2017). doi: 10.1109/TIP.2017.2713099; pmid: 28641250
83. A. N. Tikhonov, A. Goncharsky, V. Stepanov, A. G. Yagola, *Numerical Methods for the Solution of Ill-Posed Problems* (Springer, 1995), vol. 328.
84. R. Tibshirani, Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. Series B Stat. Methodol.* **58**, 267–288 (1996).
85. C. Shorten, T. M. Khoshgoftaar, A survey on image data augmentation for deep learning. *J. Big Data* **6**, 60 (2019). doi: 10.1186/s40537-019-0197-0
86. M. Prezioso *et al*., Training and operation of an integrated neuromorphic network based on metal-oxide memristors. *Nature* **521**, 61–64 (2015). doi: 10.1038/nature14441; pmid: 25951284
87. C. Li *et al*., Analogue signal and image processing with large memristor crossbars. *Nat. Electron.* **1**, 52–59 (2018). doi: 10.1038/s41928-017-0002-z
88. D. L. Andrews, M. Babiker, *The Angular Momentum of Light* (Cambridge Univ. Press, 2012).
89. L. Allen, M. W. Beijersbergen, R. J. Spreeuw, J. P. Woerdman, Orbital angular momentum of light and the transformation of Laguerre-Gaussian laser modes. *Phys. Rev. A* **45**, 8185–8189 (1992). doi: 10.1103/PhysRevA.45.8185; pmid: 9906912
90. A. M. Yao, M. J. Padgett, Orbital angular momentum: Origins, behavior and applications. *Adv. Opt. Photonics* **3**, 161–204 (2011). doi: 10.1364/AOP.3.000161
91. Z. Ji *et al*., Photocurrent detection of the orbital angular momentum of light. *Science* **368**, 763–767 (2020). doi: 10.1126/science.aba9192; pmid: 32409474
92. R. Karunasiri, Y. Mii, K. L. Wang, Tunable infrared modulator and switch using stark shift in step quantum wells. *IEEE Electron Device Lett.* **11**, 227–229 (1990). doi: 10.1109/55.55258
93. H. Jang *et al*., In-sensor optoelectronic computing using electrostatically doped silicon. *Nat. Electron.* **5**, 519–525 (2022). doi: 10.1038/s41928-022-00819-6
94. F. Zhou, Y. Chai, Near-sensor and in-sensor computing. *Nat. Electron.* **3**, 664–671 (2020). doi: 10.1038/s41928-020-00501-9
95. X. Lin *et al*., All-optical machine learning using diffractive deep neural networks. *Science* **361**, 1004–1008 (2018). doi: 10.1126/science.aat8084; pmid: 30049787
96. Y. Shen *et al*., Deep learning with coherent nanophotonic circuits. *Nat. Photonics* **11**, 441–446 (2017). doi: 10.1038/nphoton.2017.93
97. A. Bora, A. Jalal, E. Price, A. G. Dimakis, in *Proceedings of the 34th International Conference on Machine Learning* (Proceedings of Machine Learning Research, 2017), pp. 537–546.

ACKNOWLEDGMENTS

Funding: T.M. acknowledges support from the European Union (grant agreement no. 785219 Graphene Flagship). F.Z. acknowledges support from the Army Research Office under grant number W911NF-18-1-0416 and the National Science Foundation (NSF) under grant numbers DMR-1945351 through the Faculty Early Career Development Program (CAREER), DMR-2105139 through the Condensed Matter Physics (CMP) program, and DMR-1921581 through the Designing Materials to Revolutionize and Engineer our Future (DMREF) program. F.X., S.Y., and C.M. acknowledge support from the NSF Emerging Frontiers in Research and Innovation (EFRI) NewLAW program under grant number 1741693, NSF grant 2150561, the Yale Raymond John Wean Foundation, and the government of Israel. D.N. acknowledges support from the Binational Science Foundation under Electrical, Communications, and Cyber Systems (ECCS) grant NSF-BSF 2021721. **Competing interests:** S.Y., C.M., F.Z., and F.X. have a provisional patent application on intelligent light sensing. S.Y., D.N., and F.X. have a pending patent application on on-chip spectroscopy, and D.N. and F.X. have a provisional patent application on the use of on-chip spectroscopy for automobile safety monitoring. T.M. has a pending patent application on ultrafast machine vision. At present, S.Y. is affiliated with KLA Corporation in Milpitas, CA, USA; all his contributions were made based on his work at Yale University. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

Submitted 17 October 2022; accepted 15 February 2023
10.1126/science.ade1220

Downloaded from https://www.science.org at University of Texas Dallas on May 11, 2023

Geometric deep optical sensing

Shaofan Yuan, Chao Ma, Ethan Fetaya, Thomas Mueller, Doron Naveh, Fan Zhang, and Fengnian Xia

Science, **379** (6637), eade1220.
DOI: 10.1126/science.ade1220

Enhancing optical sensing and imaging
Optical sensing and imaging can be considered as an encoding/decoding process in which the encoder is the hardware or device that takes the light signal or some property thereof (e.g., intensity, polarization, or spectral composition) and transduces that signal into usable information. The decoder is the software that then takes the information and converts it into something useful for the user. Yuan *et al*. provide a review of optical sensing and imaging methods that reflect the hardware trend toward miniaturization, reconfigurability, and multifunctional ability. Simultaneously, the development of machine learning algorithms has greatly enhanced image-processing performance. The development of both areas in concert with an information theory perspective provides a powerful platform spanning many sensing applications. —ISO

View the article online
https://www.science.org/doi/10.1126/science.ade1220
Permissions
https://www.science.org/help/reprints-and-permissions

Downloaded from https://www.science.org at University of Texas Dallas on May 11, 2023

Use of this article is subject to the Terms of service

Science (ISSN) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2023 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works