



Recent advances in lensless imaging

VIVEK BOOMINATHAN,^{1,*} JACOB T. ROBINSON,¹ LAURA WALLER,² AND
ASHOK VEERARAGHAVAN¹

¹Rice University, Houston, Texas 77005, USA

²University of California, Berkeley, California 94720, USA

*Corresponding author: vivek@rice.edu

Received 1 June 2021; revised 8 November 2021; accepted 8 November 2021; published 22 December 2021

Lensless imaging provides opportunities to design imaging systems free from the constraints imposed by traditional camera architectures. Due to advances in imaging hardware, fabrication techniques, and new algorithms, researchers have recently developed lensless imaging systems that are extremely compact and lightweight or able to image higher-dimensional quantities. Here we review these recent advances and describe the design principles and their effects that one should consider when developing and using lensless imaging systems. © 2021 Optical Society of America under the terms of the OSA Open Access Publishing Agreement

<https://doi.org/10.1364/OPTICA.431361>

1. INTRODUCTION

The basic principles behind the design of a camera have not changed in several centuries. About a thousand years ago, when Iraqi scientist/philosopher Ibn-al-Haytham described an imaging device in his *Book of Optics* [1,2], he imagined it consisting of a box with some optics on one end and a sensing mechanism on another. From the earliest concrete designs for a camera, conceived by Johann Zahn [3], to their practical incarnations popularized by Niepce, Talbot, and Daugerre [4,5], cameras have remained remarkably consistent—a closed box with a focusing lens on one side and a sensor on the other. It seems no coincidence that this design resembles that of the human eye. While the last three decades have entirely transformed the technology that is used to realize both lenses (e.g., plastic molding, 3D printing) and image sensors [e.g., complementary metal–oxide–semiconductor (CMOS), charge-coupled device (CCD), single-photon avalanche diode (SPAD) detector arrays], the basic design has remained static. Lenses have served us remarkably well over the last few centuries, allowing many scientific discoveries, from the visualization of tiny bacteria to otherworldly galaxies. However, they also bring with them some limitations, especially in the context of emerging applications.

Myriad new task-based applications such as wearables, implantables, robotics, Internet of things (IoT), virtual/augmented reality, and human-computer interaction ([6–11]) are driving the miniaturization of cameras. Traditional lenses add weight and cost, are rigid, occupy volume, and have stringent focusing distance proportional to the aperture size. For these reasons, a radical redesign of camera optics is necessary to meet the miniaturization demands [12].

In a lensless camera, instead of a focusing lens, a thin, light, and potentially inexpensive optical encoder is used along with appropriate computational reconstruction algorithms to recover the

scene from captured sensor measurements. The captured sensor measurements no longer resemble the scene to be imaged—but when well designed, the measurements nevertheless contain sufficient information to recover an image of the scene. The data from these lensless imagers are usually post-processed using a computational algorithm that can demultiplex the sensor measurements and reconstruct a sharp image of the scene.

While digital post-processing of images is standard for tasks such as distortion correction [13], high dynamic range [14], synthetic depth of field (DoF) [15], noise removal [16], and low light photography [17], lensless cameras are fundamentally different in that the post-processing is a part of the imaging system design. They fall into the class of imaging systems called “computational imaging,” in which the system hardware (optics) and software (algorithms) are designed together [18]. Hence, lensless cameras encode information *indirectly* in the measurement, then computationally extract it by solving an inverse problem.

Over the last decades, several lensless imager designs have emerged, each with its own set of unique performance characteristics. Lensless imaging has been shown to have potential utility in a wide variety of applications such as microscopy [19,20], photography [21–24], *in vivo* imaging [25,26], wearables and implantables [27], and machine vision [11,28].

Advantages of lensless imaging. The principal advantages of lensless imaging are noted below.

- **Size.** Lens-based cameras, microscopes, and telescopes, by virtue of the required focusing distance between the lens and the sensor, impose strict constraints on the volume of the device. While small pixel sensors and advanced manufacturing have enabled photography with fairly small cameras on our smartphones, lensless imaging systems can reduce this volume further, resulting in ultra-compact imaging systems, sometimes reaching thicknesses of under a millimeter.

- **Weight.** Lenses account for the vast majority of the weight in conventional imaging systems. By eschewing a lens, emerging lensless systems tend to be much lighter—sometimes as light as a gram or smaller in weight.

- **Cost.** Lenses account for the largest fraction of the cost in conventional high-quality imaging systems. By removing the need for a lens, lensless imaging systems can provide an order-of-magnitude reduction in cost—sometimes requiring nothing more than an image sensor.

- **Scalability.** With the advent of lensless imaging, cameras and microscopes based on this technology can benefit tremendously from the scale-factor advantages provided by semiconductor fabrication technologies, which have so far only benefited image sensors (but not optics).

- **Field of view.** The well-known field of view (FoV) versus resolution trade-off limits conventional imaging, especially conventional microscopy. Lensless imaging provides the potential for cellular resolution imaging while achieving FoV limited only to the size of the image sensor chip, a significant advantage in some applications.

- **Visual privacy.** The captured sensor data in a lensless system is unrecognizable visually, and reconstructing an image from the data requires intimate knowledge of the parameters of the imaging system. The obscuring of the image on sensor data opens up the potential for realizing lensless cameras that provide enhanced levels of privacy while providing adequate functionality for many applications.

- **Compressive imaging.** Because lensless imagers do not focus the light from a point in the scene onto a point on the sensor, and rather map a point in the scene to many points on the sensor, they are amenable to compressive sensing approaches [29,30]. This means that one may reconstruct images with more pixels than the sensor has measurements, or higher-dimensional quantities, such as 3D imaging [25,27,31,32], hyperspectral [33], or videos from single-shot 2D measurements [34].

Limitations of lensless imaging. The principal disadvantages of lensless imaging are noted below.

- **Quality.** Lensless imaging excels in the presence of challenging constraints on weight, size, scale, and form factor. However, lenses have been the mainstay of imaging for several centuries for a reason: they produce high-quality images. Therefore, in 2D imaging applications where application-dependent constraints do not limit the use of a focusing lens, lens-based imaging should remain the preferred method.

- **Reconstruction algorithm.** All lensless imaging systems rely on computational reconstruction algorithms to recover the unknown image of the scene. This adds additional complexity to the imaging system in terms of the need for processing capability, additional power consumption, limited battery life, etc., and often means that real-time viewing of the image reconstruction is not possible or has some latency.

- **Light collection.** Often, camera lenses are designed to be larger than the sensor to collect more light, providing brighter images and better noise performance. The light collection ability of a lensless camera, on the other hand, is limited by the sensor size.

2. ANATOMY OF A LENSLESS CAMERA

A traditional lens-based camera for photography consists of a focusing lens (single or multiple elements) and an image sensor

typically placed at or near the focal distance from the lens. The lens focuses light from the scene onto the sensor, such that a point in the scene is mapped to (ideally) a single pixel on the sensor. In contrast, in a lensless imaging device, there is no lens. Instead, some kind of optical modulator (e.g., a coded amplitude mask, a diffuser) may be placed between the scene and the image sensor—typically very close to the image sensor. Then, a measurement, which looks nothing like the final image, is captured on an image sensor, and a computational inverse algorithm is solved to reconstruct the image of the scene. The relationship between the scene and the measurement can often be written as a linear matrix multiplication, so the captured image is interpreted as a weighted sum of system response functions, one from each point in the scene. Solving the inverse problem to reconstruct the image then involves inverting the system matrix, which must be known or calibrated.

Based on different optical modulations, lensless systems can be broadly classified into illumination-modulated, mask-modulated, and programmable-modulator systems (Fig. 1). This section will describe the general anatomy of a lensless camera: sensor, modulator, and illumination.

A. Sensor

For visible light sensing, conventional CMOS or CCD sensor arrays are typically used. For imaging beyond the visible domain, such as infrared imaging, sensor arrays using InGaAs sensors (for short-wave infrared) [22] or microbolometers (for long-wave infrared) [28] are available. To capture sufficient diversity of information to reconstruct an image, multiple measurements may need to be obtained. These measurements can be obtained in parallel via the use of a single focal plane array sensor with multiple pixels simultaneously capturing information or by using a single-pixel sensor but by acquiring multiple measurements with changing modulation between the measurements. Time-resolved sensors, such as SPAD [35], can also provide additional information diversity for lensless imaging [36–38].

B. Modulator

To reconstruct an image, lensless cameras need to optically encode the scene information, e.g., by having a different system response for each pixel of the scene. The system response could be a holographic response when the scene interferes with coherent light or could be a pattern produced by an optical element in the system. While attempts have been made to reconstruct images from a lensless system consisting of a sensor placed some distance from the object [39], the result is an extremely ill-posed inverse problem, since there is very little difference between measurements taken from different positions in the scene. Hence, a modulating optic (e.g., a mask placed near the sensor) is generally required for practical lensless imaging, such that the measurement is significantly different for each pixel in the scene. There are several choices for modulation (Fig. 1).

Fixed amplitude or phase mask modulators. Amplitude-only modulators are modulators that consist of transparent and occluding regions arranged in some fixed spatial pattern on a 2D mask. Most practical realizations provide only binary control of amplitude, since gray scale is generally more difficult to fabricate. Binary amplitude modulation was first used for x-ray and gamma-ray imaging [40] due to impractical lensing materials for such high-energy light. Recently, amplitude modulators have been adapted for lensless imaging in the visible, short-wave infrared, and thermal

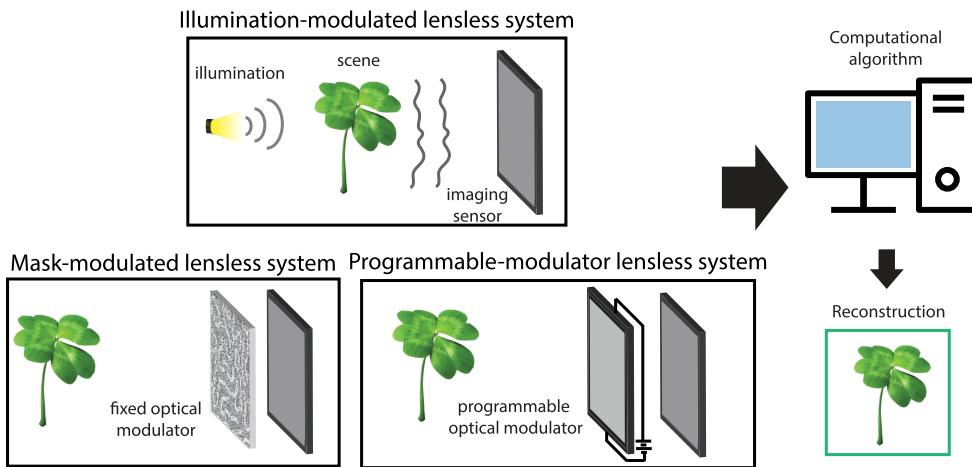


Fig. 1. Lensless cameras can be classified into several types: illumination-modulated refers to controlled illumination, mask-modulated refers to using a fixed amplitude/phase plate (modulator) in front of the sensor, and programmable-modulator systems have a programmable dynamic spatial light modulator (SLM) in front of the sensor. All types use a computational algorithm to reconstruct the scene from the captured 2D image.

wave bands. Amplitude modulation in the visible wavelengths can be achieved by etched chrome on glass [22,27,41,42] or dark ink on thin film. In general, reflective metal on glass provides better occlusion of light compared to ink. In the thermal wavelength regime, etched silicon can be used [28].

Phase modulators change the relative path length (or effective phase retardation) in a 2D pattern. Phase masks have benefits over amplitude masks in that they are more light efficient (they do not attenuate the incoming light), and they can concentrate light to create sharper and higher-contrast patterns on the sensor [23], which usually improves image reconstruction performance.

Specific designs that fall under this category of lensless imagers with a fixed mask modulator are discussed in detail in Section 4.

Programmable modulator. Dynamic and programmable modulators, also referred to as spatial light modulators (SLMs), can be used as an alternative to fixed masks. The most common SLMs are based on liquid-crystal (LC) technology [43], which is usually used for programmable amplitude modulation [44–46]. For realizing programmable phase modulation, LC on silica (LCoS) devices may be used [47].

The primary benefit of a programmable modulator is that one can change the mask pattern quickly and capture multiple images, each with different optical encoding. In imaging designs with limited numbers of pixel photodiodes [44,48,49], it is essential to use programmable modulators to obtain sufficient measurements for reconstruction. When sensor arrays are used, the ability to change the modulation pattern between acquisitions provides an extra degree of freedom that can improve reconstruction performance or resolution. The main drawback of a programmable modulator is that when multiple measurements are needed for reconstruction, the image capture time increases. Additionally, LC-based technologies are polarization sensitive, and diffraction from the pixel grid often causes unwanted light scattering. Specific designs that fall under the category of lensless imagers with a programmable modulator are discussed in detail in Section 5.

Illumination modulators. Some lensless designs instead use illumination devices to pattern the light in the system, rather than detection-side patterning. Depending upon the property of the illumination used, the techniques are: shadow imaging, holographic imaging, or time-resolved imaging. Specific designs

that fall under the category of lensless imagers with illumination modulator are discussed in detail in Section 3.

C. Illumination

Illumination is a critical part of any imaging system. Below, we highlight the most salient illumination strategies employed.

Ambient illumination. In some imaging scenarios, especially in macro-imaging, consumer imaging, or photography contexts wherein subjects are far away from the camera, ambient illumination scattering from subjects is sufficient, and no additional considerations are necessary.

Trans-illumination. When access to both sides of a sample is available, trans-illumination can be used to illuminate the sample from one side and image from the other side. Many microscopes, including lab-on-a-chip applications, use this configuration with the sample placed on a glass slide and inserted between the illumination and the imaging sides of the imager.

Off-chip, side illumination. In many *in vivo* microscopic imaging applications, ambient illumination is insufficient, and trans-illumination is physically impossible. Therefore, one needs to actively illuminate the sample from the same side as the imager (epi-illumination). One solution for this is to have one or more LEDs placed around the sides of the lensless imager directing light towards the sample [26]. For microscopy, where the sample is placed very close to the imager, uniformly illuminating the sample is challenging and requires careful design with a combination of optics (optical fiber, prisms, etc.) and mechanical placement of the sources (raised platform, angled-in, etc.).

On-chip, integrated illumination. When the FoV being imaged is very large and the sample is close to the imager, the ideal configuration is an on-chip array of light sources, such as light-emitting diodes (LEDs) or vertical-cavity surface-emitting lasers (VCSELs), for illuminating the sample [50]. However, such a design introduces a whole host of new challenges that need to be addressed, in addition to the fabrication challenges of such a design. First, the on-chip sources occupy the space that would have otherwise been image sensor pixels—meaning that the acquired lensless sensor data have missing pixels that need to be interpolated or otherwise accounted for in computational reconstruction.

Second, light leakage directly to the sensor pixels from these on-chip sources becomes a problem, especially for the pixels near the sources.

Coherent illumination. Digital holography-based lensless systems rely on the spatial and temporal coherence of light sources. In such cases, laser diodes or narrowband LEDs can be used. Given that multiple previous review papers [20,51] have focused on this large area of research, we limit our scope primarily to non-holographic lensless imagers.

D. Sensor-Mask Distance

A primary advantage of mask-modulated lensless cameras is that the sensor-to-mask distance can be extremely small, as compared to the sensor-to-lens distance in a lens-based camera. This allows the overall thickness of a lensless system to be extremely small (less than 200 μm [27,52]) while maintaining a large FoV. The mask design dictates the ideal sensor–mask distance; masks with smaller features will have a smaller sensor–mask distance and larger angular FoV, but worse resolution, as compared to masks with larger features and larger sensor–mask distances. Larger features are usually easier to fabricate and more tolerant to misalignment, and when the sensor–mask distance is of the order of millimeters, one can conveniently affix the mask to a commercial sensor without removing the protective coverglass [11,23,53]. On the other hand, masks with very small features can be fabricated very precisely on-chip with the sensor for a compact and stable design [27,50]. Hence, the sensor–mask distance is a design degree of freedom that can be optimized based on the application.

E. Object Distance

The working distance from the lensless imager is often not a design degree of freedom but rather an application-dependent constraint, which has significant implications for the design in two ways. First, the object distance imposes physical limits on what illumination strategies can be incorporated. As the object distance becomes larger, there are fewer physical constraints on illumination devices, and the smaller numerical aperture (NA) gives more flexibility for achieving near-uniform illumination of the target. In photography, for example, ambient illumination is usually sufficient, whereas in fluorescence microscopy, one must design for both the illumination and detection devices. Second, the absence of a focusing lens means there is no magnification in lensless imaging devices; the system is non-telecentric, and resolution varies with depth. Hence, spatial resolution δs is best characterized as angular resolution $\delta\alpha$, which can be computed from the target distance d as $\delta s = d * \tan(\delta\alpha)$. One important aspect to remember here is that many applications have a desired spatial resolution (e.g., 5 μm to image cells) that, given a fixed angular resolution, sets a maximum distance the target can be from the sensor.

3. ILLUMINATION-MODULATED LENSLESS SYSTEM

Illumination-modulated lensless systems rely on how controlled illumination interacts with the sample. Properties of illumination such as location, coherence, and pulsing are exploited to capture a set of images with different illumination diversity, then reconstruct the scene.

A. Shadow Imaging

The simplest form of illumination-modulated lensless imaging uses a constant illumination source and an image sensor. Commonly used for lab-on-a-chip microscopy applications, the illumination and the image sensor are placed on opposite sides of the microscopic sample (Fig. 2A), usually with a larger illumination-to-sample distance than the sample-to-sensor distance. When the sample is very close to the sensor, the result is a shadow image [54]; when the sample is slightly farther away, such that it diffracts (defocuses) slightly, the result is diffraction imaging [55]. Diffraction-based techniques can handle larger sample–sensor distances compared to shadow imaging and are discussed in the next section. The shadow imaging scheme requires samples to be placed as close as possible (typically less than 500 μm) to the surface of the imaging sensor [20].

Lange *et al.* [54] demonstrated a miniaturized microfluidic shadow imaging device for studying *C.elegans*. This early system was very compact, but the low resolution ($>10 \mu\text{m}$) limited its applications. Ozcan *et al.* [19] demonstrated a wide FoV imaging platform that was two orders of magnitude larger (37.25 mm \times 25.70 mm) than the conventional optical microscope for monitoring the cells on-chip, termed LUCAS. This system recorded the shadow image of cells onto the sensor plane to monitor and count different cell types. Because of the large pixel size sensor (9 μm) used in the system, its spatial resolution was limited.

Pixel size limitations can be overcome by flowing the sample over the sensor and exploiting the temporal dimension to improve resolution. To physically sample a higher resolution than pixel size, Heng *et al.* [56] and Cui *et al.* [57] fabricated a linear array of sub-micrometer metallic apertures on the sensor with the apertures centered on the pixels. By flowing the sample, *C.elegans*, in this case, within a microfluidic channel at a titled angle over the linear array, adjacent line scans of the sample are projected onto the pixels. By estimating the flow velocity, the line scans can be un-skewed to create an image of sub-micrometer resolution. The resolution of this optofluidic microscope is fundamentally limited by the size of the apertures and the spacing in between. Zheng *et al.* [58], on the other hand, opted to use computational methods instead of physical apertures to reconstruct sub-micrometer resolution images from a sequence of pixel-size-limited images. As the sample flows over the sensor, sub-pixel shifted projections are captured as a sequence of low-resolution images. A pixel super-resolution algorithm can be used to combine the low-resolution images to produce a single high-resolution image (Fig. 2B).

B. Holographic Lensless Imaging

A lensless holographic imaging system uses a coherent or partially coherent light source to obtain a diffraction pattern on the sensor [59,60]. Similar to shadow imaging, lensless holographic imaging is commonly used for lab-on-a-chip microscopy applications and has the microscopic sample placed in between the light source and the sensor (Fig. 2A). However, since holographic imaging encodes complex-field information (amplitude and phase) instead of just amplitude information, the sample can be reconstructed at higher fidelity than shadow imaging. The coherence of the light source is a critical parameter in holographic imaging; lasers provide good spatial and temporal coherence, but can suffer from coherent speckle noise. LEDs can be used to tune the spatial coherence according to

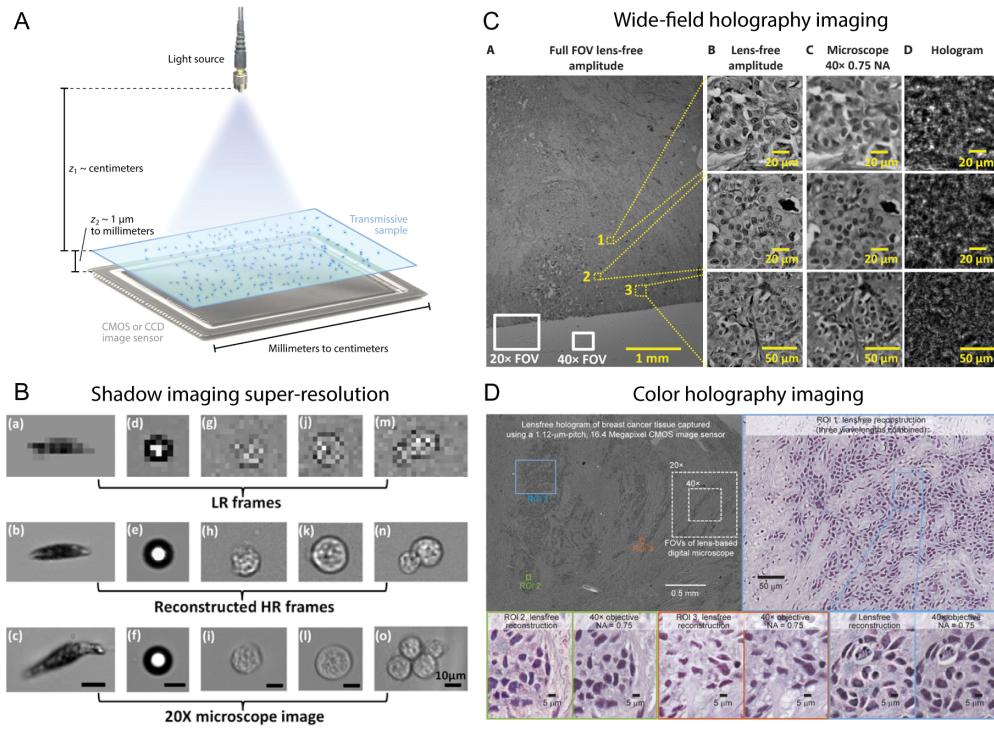


Fig. 2. Illumination-modulated lensless microscopy systems. (A) In this case, the sample is placed close to the sensor and trans-illuminated by a light source placed much farther from the sample (modified from [20]). (B) Shadow imaging super-resolution by combining multiple low-resolution frames with sub-pixel shifts (image modified from [58]). (C) Wide field of view holographic imaging [67]. (D) Color holographic imaging by combining reconstruction from three wavelengths of illumination (image modified from [68]). (C) and (D) show the captured hologram, reconstructed lensless image, and comparison with the image taken with a high-NA microscope.

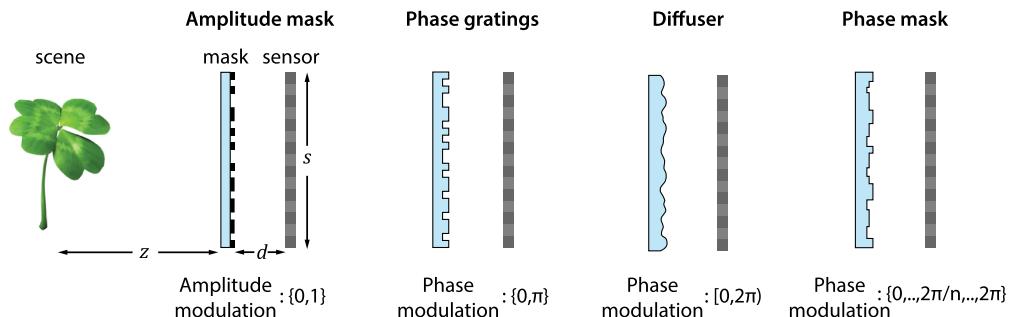


Fig. 3. Types of masks used in mask-modulated lensless imaging systems. Binary amplitude masks either block or allow light to pass through, giving amplitude modulation of either zero or one. Binary phase gratings have two heights of transparent material, giving phase modulation of either zero or π . Diffusers have continuous but random surface heights for continuous phase modulation. Multi-level phase masks have “ n ” discrete material heights, giving discrete phase modulations.

the size of the source, its bandwidth, and its distance to the sample [61,62].

An approach to reconstruct is to consider the diffraction pattern on the sensor as an **in-line hologram** [63,64]. An in-line hologram is the intensity pattern generated by the interference between the light scattered by an object on the sample and a reference wave that passes undisturbed through the sample. Then, the sample can be reconstructed from the measurements by digitally backpropagating using the Fresnel diffraction integral [65]. An in-line hologram reconstruction method was used to show high-resolution images of red blood cells on lab-on-chip devices [66]. However, this basic implementation of digital holographic reconstruction is limited to only relatively sparse samples because the scattered waves need to be weak compared to the reference wave.

A more general approach is to consider the holographic measurement as a coherent diffraction pattern without any assumption of a clean reference wave. Under this general consideration, additional information such as prior knowledge of object support in the sample (such as sparsity) [62] or multiple measurements under different device geometries [67–70] are needed for recovery of the sample. Iterative phase retrieval algorithms [71] are used to incorporate the additional information and reconstruct images of the sample. Multiple measurements, in particular, can be used to reconstruct dense samples such as tissue slices or blood vessels. Greenbaum *et al.* [67] captured multi-height measurements by varying sample-to-sensor distance to reconstruct a human breast cancer tissue slice (Fig. 2C). Luo *et al.* [68], instead, captured

multiple measurements under different angular illuminations to reconstruct the human breast cancer tissue slice (Fig. 2D).

Holographic lensless imaging has been a focus for previous review papers [20,51], and we refer the reader to those papers for further details.

C. Time-Resolved Lensless Imaging

Pulsed illumination coupled with time-resolved sensors can use time-of-light information to image the scene [72]. High time resolution (order of picoseconds [73]) of such systems has enabled its success in the challenging case of non-line-of-sight imaging around a corner [74,75]. Similar concepts have been used to demonstrate lensless imaging; however, the extent of work in this direction is limited. Kirmani *et al.* [36] showed, through simulation, lensless imaging with a pulsed light source and a small array time-resolved sensor. Wu *et al.* [37] showed real experiment lensless imaging but with an expensive and bulky streak camera [76]. Satat *et al.* [38] proposed a time-resolved lensless imaging framework with spatial illumination patterning, sensor placement optimization, and using compressive sensing principles to reconstruct. However, the results were limited to numerical simulations.

4. MASK-MODULATED LENSLESS SYSTEM

Within the class of mask-modulated lensless systems, a fixed optical mask is introduced to make a versatile lensless system that can work for a large range of object distances and passive or uncontrolled lighting scenarios. The mask modulates the incoming light to produce an encoding that can be decoded using computational techniques. Mask-modulated lensless cameras were used to demonstrate 2D imaging [22–24], refocusing [23], 3D imaging [23,53], and microscopic imaging [23,25,27]. Example imaging results from mask-modulated lensless imagers are shown in Fig. 5. The majority of recent lensless designs fall into this category.

A mask-modulated lensless camera consists of an optical mask placed in front of an image sensor, often imaging a scene with ambient illumination. The mask modulates the light from the scene and forms a multiplexed measurement on the sensor. For the

typical mask-modulated lensless system shown in Fig. 3, we can say the following.

- **Magnification.** For a given mask-to-sensor distance, the magnification M is

$$M = \frac{d}{z} \quad (1)$$

where d is the mask-to-sensor distance, and z is the scene-to-mask distance. This equation is similar to that of a fixed focus simple lens system, where d would be the image distance and z the object distance.

- **FoV.** The angular FoV is given by

$$\theta_{\text{FoV}} = \min(2\theta_{\text{CRA}}, 2 \tan^{-1}(s/(2d))) \quad (2)$$

where θ_{CRA} is the chief ray angle (CRA) of the pixels, and s is the sensor size. When d is small, the FoV is limited by the CRA [22]. When d is large, the sensor size limits the FoV [46]. The CRA describes the maximum angle of incoming light for which the pixels are fairly sensitive. That is, light entering the pixel from angles larger than $\pm\theta_{\text{CRA}}$ with respect to its surface normal will be captured with very low sensitivity.

The optical masks used for lensless cameras can be broadly categorized into amplitude modulators [21,22,27,77] and phase modulators. Phase modulators can be further sub-categorized into phase gratings [78,79], diffusers [53,80], and phase masks [23]. Illustrations of these masks are shown in Fig. 3. We will go through each below.

One key characteristic of a mask-modulated lensless system is the pattern that the mask produces on the sensor for a point source in the scene. We call this pattern the point-spread function (PSF), and its properties determine the imaging model of the system (Section 6.A). We will go through the various choices of PSFs after we describe the main types of optical masks.

A. Amplitude Modulators

An amplitude mask passes, blocks, or attenuates incident light. For ease of fabrication, binary amplitude masks are most commonly

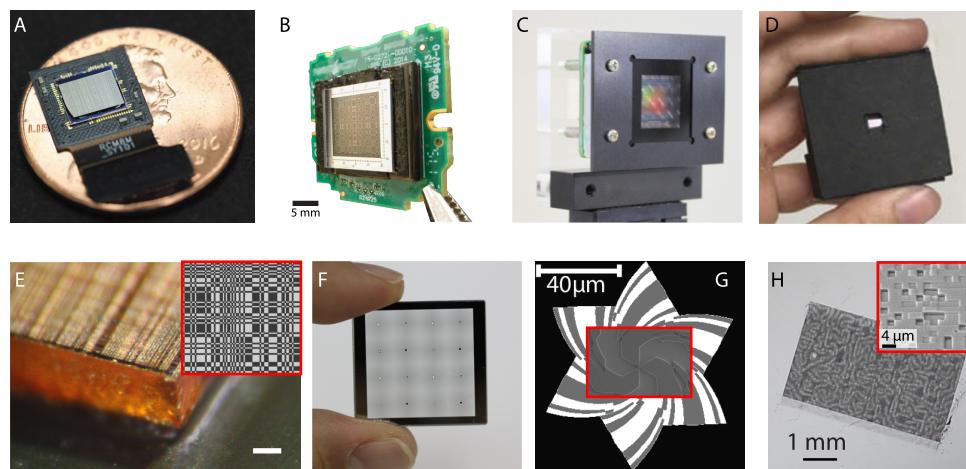


Fig. 4. (Top) Images of various lensless camera prototypes. (A) FlatScope (image modified from [27]). (B) FlatCam [22] (image modified from [11]). (C) Fresnel zone aperture (FZA)-based lensless imager (image modified from [77]). (D) Random microlens diffuser microscope (image modified from [25]). (Bottom) Images of various masks used. (E) Separable amplitude mask (image modified from [27]). (F) Amplitude mask with multiple FZAs of different fringe phases (image modified from [77]). (G) Spiral phase gratings (image modified from [79]). (H) Phase mask that generates contour PSF (image modified from [23]).

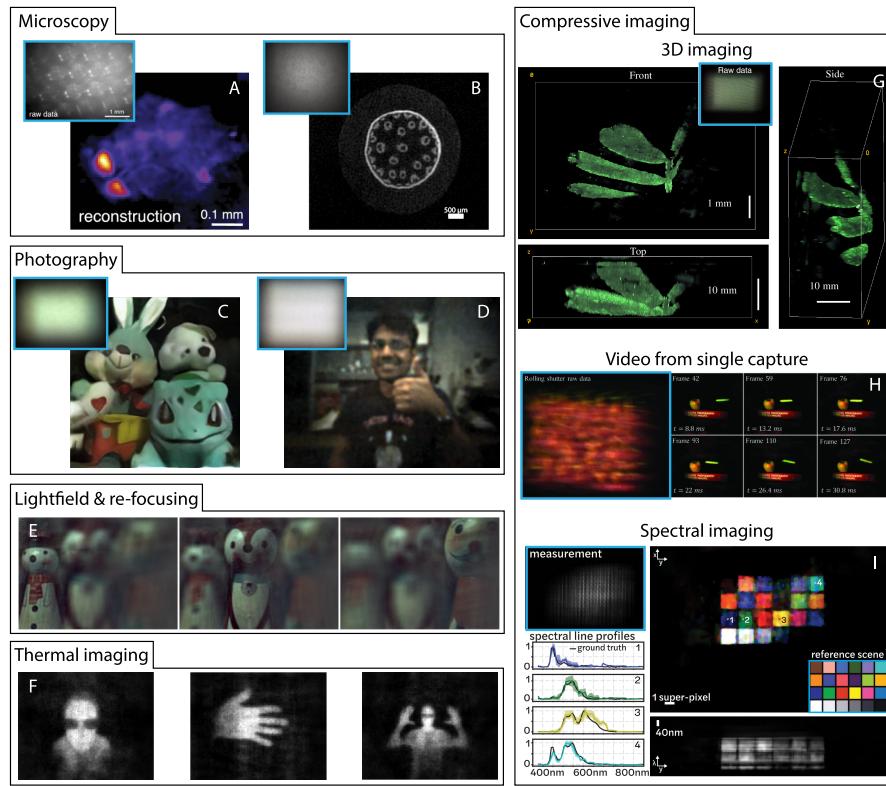


Fig. 5. Example mask-modulated lensless imaging results. Microscopy: (A) modified from [25], (B) modified from [23]. Photography: (C) modified from [103], (D) modified from [23]. (E) Light field and re-focusing, modified from [46]. (F) Thermal imaging, modified from [28]. Compressive imaging: (G) 3D image, modified from [53], (H) video from single capture, modified from [34], and (I) spectral imaging from single capture, modified from [33].

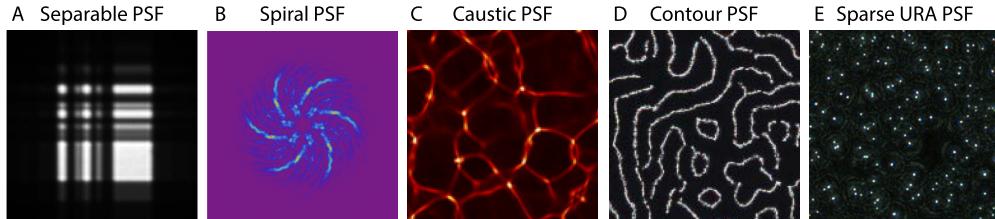


Fig. 6. Experimental point-spread functions (PSFs) of mask-modulated lensless systems. (A) Separable PSF, generated by the amplitude mask in [21]. (B) Spiral PSF of binary phase gratings, used in [79]. (C) Caustic PSF of diffuser used in DiffuserCam [53]. (D) Contour PSF of phase mask used in PhlatCam [23]. (E) Sparse URA PSF, by amplitude mask, used in [84]. All images were modified from respective references.

used. The simplest binary amplitude mask is a pinhole; however, the diminutive light throughput makes it extremely impractical. Early work on amplitude mask modulators with multiple “pinholes” or apertures was in the x-ray and gamma-ray imaging regimes [40,81], where it is much easier to block than refract high-energy light. Some of the lensless designs in the visible wavelengths were inspired by the early work in x-ray and gamma-ray imaging [12].

Light modulation by amplitude masks is achieved and modeled in two ways: **as a shadow of the mask when mask–sensor distance is small** and **as diffraction when mask–sensor distance is large**. The distinction can be roughly based on the Fresnel number $N_F = a^2/d\lambda$ [65], associated with the size a of the mask’s open apertures, distance d to the sensor, and wavelength λ of light. If N_F is much greater than one, then geometrical properties are valid, and the PSF mimics the shadow of the mask pattern. On the

other hand, when N_F is less than one, diffraction effects need to be considered to determine the PSF.

Asif *et al.* [22] used an amplitude mask with a separable pattern for their FlatCam device and placed it close (1.2 mm [12]) to the sensor (Fig. 4B). The PSF of FlatCam was modeled as the shadow of the pattern. A separable pattern is constructed as an outer product of two 1D patterns, and using such a pattern for the PSF simplifies the computational model of the lensless camera (see Section 6.A). FlatCam showed the promise of using amplitude masks for creating very thin cameras [12]. Adams *et al.* [27] further miniaturized this design (200 μm mask–sensor distance) with their FlatScope device (Figs. 4A and 4E), to show applications in 3D fluorescence microscopy. FlatScope showed the promise of breaking the trade-off between FoV and resolution, placed by microscopy lenses, by experimentally showing an FoV of 6.52 mm^2 , 10 \times the FoV of a microscope with a similar resolution objective lens and the same sensor size.

Nakamura, Tajima, and Shimano *et al.* [41,77,82] introduced a Fresnel zone aperture (FZA) as the amplitude mask pattern placed close (2 mm) to the sensor (Figs. 4C and 4F). At this distance, the PSF is the shadow of the FZA pattern itself. Ghosting effects in the image reconstruction are subdued by combining measurements through a set of four to 16 radially-phase-shifted FZA patterns. Wu *et al.* [42] showed imaging with a single FZA mask (3 mm mask–sensor distance), using a compressive sensing algorithm [83] to reconstruct.

When the distance of the amplitude mask to the sensor increases to more than tens of millimeters and with smaller mask feature sizes, the Fresnel number becomes less than one, and diffraction effects become significant. DeWeert *et al.* [21] used a separable amplitude pattern (at 65 mm from the sensor) whose PSF remained separable even with diffraction effects. A separable Doubly-Toeplitz model was then used to reconstruct cityscape images. Reshetouski *et al.* [28] used Fresnel zone plates (FZP) to replace pinholes and achieve relatively light-efficient focus points within the PSF. With thermal wavelengths, short mask–sensor distance (3.2 mm [28]) is sufficient for significant diffraction ($N_F < 1$). In the visible wavelengths, Reshetouski *et al.* [84] achieved the same diffraction focus points by placing the FZP mask at a longer distance of 38.6 mm from the sensor.

Amplitude masks generally have the advantage of being easier to fabricate for a wide range of wavelengths. Outside the visible range, materials that can block light are easier and cheaper to find than those that can refract light. Chromium (Cr) is the material of choice for blocking light in the visible and thermal wavelength ranges, because of its good adhesion to glass. Cr is deposited on a transparent wafer (glass for visible light, silicon for thermal), patterned using a photo-lithography process and then etched to produce the mask pattern [27,42].

A major disadvantage of amplitude masks is the light throughput. Since the mask modulates light by either passing or blocking light, many photons are lost, leading to low signal-to-noise-ratio (SNR) measurements. Low SNR is particularly problematic in low-light scenarios and photon-limited imaging such as fluorescence or bioluminescence imaging. Decoding the lensless sensor capture tends to amplify noise, leading to poor reconstruction.

B. Phase Modulators

A phase modulator modulates the phase of incident light by the principles of wave optics [85]. Phase modulators allow almost all of the light to pass through, providing high SNR. Hence, they are desirable for low-light scenarios and photon-limited imaging.

Phase modulators can be sub-categorized into phase gratings, diffusers, and phase masks. Phase gratings and masks can be fabricated by either by photo-lithographically etching patterns into a transparent substrate (e.g., glass) or by an additive process of controlled polymerization of photoresist on the transparent substrate [86]. Diffusers can be made by etching random surface heights onto glass or inexpensive sandblasting to produce a rough surface on the glass.

1. Phase Gratings

Phase gratings modulate the phase of incoming light at each lateral pixel by 0 or π rad. Stork and Gill [79] exploited an odd-symmetry binary phase structure (Fig. 4G) to produce nulls or zero intensity regions on the sensor. The gratings were used to demonstrate

ultra-miniature lensless imaging [52,78] and thermal lensless imaging [87]. However, the imaging performance of these systems was limited to low resolution.

2. Diffusers

Diffusers are an inexpensive way to achieve phase modulation. Diffusers have continuous random height profiles and produce pseudorandom patterns as the PSF. Singh *et al.* [80] introduced scatter-plate microscopy, which used the speckle pattern produced by a diffuser as the PSF. The auto-correlation of this speckle pattern is close to a “ δ ” function with a constant background so that the sample image can be reconstructed by cross-correlating the measurements with the PSF. Antipa and Kuo *et al.* [24,53] introduced DiffuserCam, in which a diffuser (with Gaussian 0.5° FWHM angular spread) was placed ~9 mm in front of the sensor, generating a high-contrast caustic pattern PSF (Fig. 6C). Using convex optimization techniques for image reconstruction with compressed sensing, DiffuserCam was demonstrated for 2D [24] and 3D imaging [53]. For microscopy, Kuo *et al.* [25] fabricated a random microlens diffuser (placed at 3.8 mm mask–sensor distance, Fig. 4D) to improve the contrast of the PSF and achieved improved reconstruction quality in low-light scenarios of fluorescence microscopy. Tian *et al.* [88] designed a random microlens array with reduced inter-lens cross talk to simplify the image reconstruction process and drastically improve reconstruction speed.

3. Phase Masks

Phase masks provide nearly continuous phase modulation, limited only by the fabrication method used. With such control, phase masks can achieve a large range of PSFs. However, relative to previous phase modulators, fabricating custom phase masks is more involved and can be relatively expensive. Boominathan *et al.* [23] introduced PhlatCam with a designed phase mask (Fig. 4H) that produced a very high contrast contour pattern as the PSF. Inspired by Chi *et al.* [89,90], PhlatCam used phase retrieval methods [91] to design the phase mask for the target contour PSF. The versatile PhlatCam (~2 mm mask–sensor distance) prototype was used for 2D imaging at far distances, refocusing at medium distances, 3D imaging at close distances, and microscopy at very close sample–mask distance. Adams *et al.* [26], used the mask design from PhlatCam and rebuilt the system with integrated illumination and wavelength filter, called FlatScope2.0, for *in vivo* microscopy applications.

C. Point-Spread Functions

All mask-modulated systems can be characterized by their response when illuminated by a single point source of light. This response is called the PSF, whose properties are critical for image recovery. Generally, the desired properties of a lensless PSF is that its auto-correlation is close to a “ δ ”-function and has an almost flat magnitude spectrum. This allows for maximal information transfer onto the measurements and reliable recovery of images [23].

The PSF design, in some cases, is coupled with the choice of mask used. Amplitude masks, in most cases, are used to produce binary PSFs, but the PSF could deviate from the target PSF due to diffraction [21,23]. Phase gratings are used to produce reliable

nulls in the PSF by introducing an odd-symmetric phase [52,78]. Diffusers produce caustic patterns as their PSF (Fig. 6C) when the mask–sensor distance is at the caustic plane [24,53] and speckle pattern PSFs [80] at farther distances. Phase masks were shown to achieve a larger range of PSFs [23]; however, the achievable space of the PSFs is less understood due to the non-convex nature of the phase mask design using phase retrieval methods.

A plethora of PSFs have been explored for use in mask-modulated lensless imaging, many of which are realized using amplitude masks. A separable PSF (Fig. 6A) was shown to reduce the computation complexity of the camera model to two smaller operations (see Section 6.A for more details), each applied on the rows and columns of the image [21,22,27]. A separable PSF, by definition, is constructed by an outer product of two 1D vectors. Patterns based on a uniformly redundant array (URA) [92], modified URA (MURA) [93], and maximum length sequence (MLS) [94] are popular choices due to their auto-correlation properties.

An FZA PSF, constructed like an FZP [95] amplitude mask, was used by [41,77]. When the measurements are post-multiplied by the same FZA pattern, it was observed that the scene spatial information was transferred to the magnitude in the frequency domain through the resultant moiré fringes. This led to a one-step analytical solution to reconstruct images from the Fourier transform of the moiré fringes. However, the frequency domain also contains information that interferes with reconstruction. Therefore, the scene is imaged through a set of four to 16 phase-shifted FZA PSFs and then combined to subdue this unwanted information. A single FZA PSF was also used where reconstruction was done with iteratively solving a convex minimization problem [42]. On the other hand, FZP patterns were used for its original intended property of diffractively focusing light to a point [95] to create a sparse URA PSF (Fig. 6E) [28,84].

By controlling the nulls in the PSF using odd-symmetric bi-level phase gratings, a spiral PSF (Fig. 6B) was proposed [52,78]. A tessellation of spiral PSFs was proposed to cover a large sensor area [79]. However, precise control of the PSF is difficult with this design. On the other hand, a multi-level phase mask can be designed to achieve a wide variety of PSFs. A contour PSF [23] (Fig. 6D) was proposed to have the desirable properties of an invertible PSF and was realized using a phase mask designed using phase retrieval methods [71].

5. PROGRAMMABLE MODULATOR LENSLESS SYSTEM

Various SLM technologies such as LCD, LCoS, and digital micromirror device (DMD) can provide dynamic programmable optical encoding [44–46]. Programmable modulators allow for multiple lensless images to be captured for each reconstructed image; the different diversity in the measurements then leads to better reconstruction [21] or allows very few pixels on the sensor [44]. Multiple measurements can also be used to simplify the reconstruction algorithm while maintaining high fidelity [46].

Huang and Jiang *et al.* [44,96] used an LCD panel as a transmissive amplitude modulator to compressively [29,30] capture scene images onto just one or two pixels. The compression comes from the fact that the number of measurements taken is a fraction (12.5% to 25%) of the number of pixels in the final reconstruction image. Expansion up to a 4×4 pixel array, with spaced apart pixels, was explored to split the FoV [97]. Zomet and Nayar [45] used

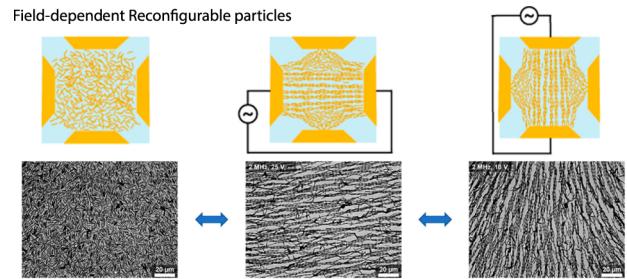


Fig. 7. Reconfigurable nanoparticles to achieve dynamic scattering masks [98]. The particles are silica-coated gold nanowires suspended in water to form a scattering mask. The nanowires are highly polarizable, exhibit a strong response to the applied electric field, and can be oriented in different ways depending on the direction of the applied field. Image modified from [98].

an LCD panel to act as a controllable pinhole to track a moving object onto a full-resolution sensor, without any moving parts.

Multiple measurements through different mask patterns capture the scene onto the sensor through different “filters.” Each filter may capture some part of the scene information reliably while discarding other information. Such information differences could be characterized in the Fourier domain or similar analysis, and combining the different filtering can lead to a better reconstruction. Although [21] showed reconstruction with single-shot amplitude modulation, the artifacts in the reconstruction were substantially reduced by incorporating two to four measurements with different mask patterns. Via simulation, [12] showed that increasing the number of measurements improved the reconstruction asymptotically, with higher gains in the beginning and diminishing returns at a larger number of measurements. Additionally, there is a trade-off when trying to image dynamic objects, which time-limits the number of images that may be captured for each reconstruction.

Miller *et al.* [98] introduced a different way of achieving diversity in measurements by using reconfigurable particles as scattering masks (Fig. 7). The particles are silica-coated gold nanowires suspended in water to form a scattering mask. The nanowires are highly polarizable, exhibit a strong response to the applied electric field, and can be oriented in different ways depending on the direction of the applied field. This technology is similar to LC technology, except that multiple random particle configurations are achieved here. The final reconstruction was improved by measuring with and without applied field and under different directions of the field. As before, the improvement in reconstruction is asymptotic with the increase in the number of measurements.

Hua *et al.* [46] used a translating mask, implemented using LCoS, to increase the diversity in terms of the viewpoint of the camera. This leads to a depth-dependent translation of measurements, and the information from a particular depth can be enhanced by summing the appropriately shifted measurements. Reconstructing from such combined measurements produces an image focused at a particular depth plane. By designing the amount of shift and the number of shifts, the effect of the focusing operation can be made to depend minimally on all depths, except one. This allows decoupling the optimization problem of joint texture recovery of all depths to just individually solving each depth plane, resulting in a fast recovery.

6. IMAGE RECONSTRUCTION

The absence of a lens implies that the measurements in a lensless camera are not direct measurements of the scene intensity. Instead, the particular design of the lensless imaging device induces a relationship (in most cases, this relationship is linear) between the sensor measurements and scene intensity that needs to be recovered by computational reconstruction.

A. Forward Model

The relationship between the sensor measurements and scene intensity is typically represented using an appropriate forward model that calculates the expected measurement for a given scene. Often, the optical physics involved is complicated and difficult or computationally intensive to model comprehensively. The main goal of the forward model is to choose the right approximations for this relationship (typically related to the various assumptions that one can make about the scene being imaged) that achieve a practically useful trade-off between accuracy and reconstruction complexity.

1. Forward Model for 2D Imaging

The simplest case for the forward model arises when the scene can be assumed to be a single planar target with 2D spatial intensity variations in the target plane to be imaged. In this case, one can usually write the relationship between the captured sensor measurements \mathbf{y} , and the unknown scene intensity \mathbf{x} using linear matrix multiplication as

$$\mathbf{y} = \mathbf{H}\mathbf{x}, \quad (3)$$

where \mathbf{H} is a matrix that is dependent on the various design choices made. In particular, the elements of the matrix \mathbf{H} are dependent on the choice of sensor, choice of mask, sensor–mask distance, mask–scene distance, and any other system design parameters. While this general forward model is broadly applicable for a wide variety of lensless imaging system designs, in the absence of any further assumptions it will be challenging to implement practically for sensors with large pixel counts. Consider, for example, a megapixel sensor with a megapixel image being reconstructed: \mathbf{H} is a $10^6 \times 10^6$ matrix, so there are 10^{12} elements in the matrix, and every application of the forward model requires 10^{12} multiplications, making it infeasible in terms of both memory complexity and computational complexity. To significantly reduce the computational complexity of the forward model, two ideas have emerged as powerful alternatives: convolutional and separable approximations.

Convolutional model. Under certain conditions such as narrow FoV, sufficient mask–sensor distance, and far-field approximation, one can show that the general linear model can be approximated as a convolution with a PSF (i.e., the relationship between scene intensity and sensor measurements is shift invariant) [23,53]. In this case, the relationship between the captured sensor measurements $\mathbf{Y}(x, y)$ and the unknown scene intensity field $\mathbf{X}(x, y)$ can be written as

$$\mathbf{Y} = \mathbf{X} * \mathbf{h}, \quad (4)$$

where $\mathbf{h}(x, y)$ is a 2D kernel (the PSF), and $*$ denotes 2D convolution over (x, y) . The benefit of this convolutional approximation is two-fold. First, instead of saving the entire $10^6 \times 10^6$ matrix \mathbf{H} ,

one needs to save only a single 2D kernel \mathbf{h} . Second, the forward model can be efficiently implemented in the Fourier domain, which reduces computational complexity from $\mathcal{O}(N^2)$ to $\mathcal{O}(N \log N)$, where N is the number of pixels—resulting practically in three orders-of-magnitude reduction in computational complexity. A cropped convolution model [53] (convolution followed by cropping) can be used when the sensor is not large enough to contain all the light coming from wide FoV angles.

The main limitation of the convolutional model is that when the FoV increases or equivalently when the object-to-mask distance significantly reduces (for applications such as microscopy), the shift-invariance assumption breaks down—and the convolutional model is no longer a good approximation. One technique that has been explored to extend the applicability of this model is a local convolutional model [25], wherein the convolution kernel is assumed to change slowly over the FoV.

Separability model. Another alternative for reducing the computational and memory complexity of the general linear model is through separability. A mask whose 2D spatial pattern is separable, i.e., can be represented as a cross-product of two 1D functions, results in a separable relationship between the captured sensor measurements and scene intensity [21,22]. In this case, the relationship can be compactly represented as

$$\mathbf{Y} = \Phi_L \mathbf{X} \Phi_R^T, \quad (5)$$

where Φ_L and Φ_R represent the separable operations on the image's rows and columns, respectively. As such, the main advantage of this separability is that the Φ_L and Φ_R matrices are much smaller than \mathbf{H} in Eq. (3). For the example of a megapixel sensor and image, the separable matrices are $10^3 \times 10^3$, reducing the memory needed by six orders of magnitude. In addition, since the separable matrix operation is massively parallel and can be applied row by row and then column by column, this also reduces the computational complexity to $\mathcal{O}(N)$, where N is the number of pixels. An extended separable model called the Texas Two-step Model [27] allowed the use of large separable pattern masks for compact and wide FoV microscopy applications.

2. Forward Model for 3D Imaging

All of the models developed above can be extended to 3D imaging. The key aspects to remember are: (a) the transfer functions that relate scene-intensity-to-sensor measurements are depth dependent, and (b) light distributions arriving from multiple depth planes are additive at the sensor. Given this, the general linear model, the convolutional model, and the separable model can all be extended to 3D as

$$\mathbf{y} = \sum_z \mathbf{H}_z \mathbf{x}_z, \quad \mathbf{Y} = \sum_z \mathbf{X}_z * \mathbf{h}_z, \quad \mathbf{Y} = \sum_z \Phi_{Lz} \mathbf{X}_z \Phi_{Rz}^T, \quad (6)$$

where the subscript z represents the depth dependence of scene features or transfer functions, as appropriate, and \sum_z denotes a sum over the z dimension. The computational and memory complexity of these forward models increases linearly with the number of depth layers used in the model, but is otherwise very similar to the 2D models above in terms of their limitations.

3. Light Field Representations

An alternative representation that allows for post-capture extraction of 3D and viewpoint-dependent information is the light

field parameterization [99,100]. In this representation, instead of recovering 2D or 3D intensity distributions, we attempt to recover 4D light radiance $\mathbf{L}(x, y, \theta, \phi)$ of rays intersecting a plane at (x, y) and traveling in the direction (θ, ϕ) . This 4D representation is redundant and therefore has more parameters to estimate compared to the 3D intensity field, but forward models and reconstruction algorithms can efficiently exploit the elegant geometry that this representation affords. All previously developed forward models (general linear, convolutional, and separable) can be adapted to the light field representation.

4. Holographic Representation

A highly general approach to extracting 3D spatial information contained within the incoming light wavefront is to represent the wavefront as spatially coherent, using a complex-field representation. This model is valid only in holographic imaging and other scenarios where light is sufficiently coherent.

5. Video Imaging Models

The discussion so far has considered only an instant in time; however, dynamics captured in multiple frames are often of interest. For video imaging, the dynamics may be decoupled, so both the forward model and reconstruction algorithms are implemented on a frame-by-frame basis, assuming that any motion and dynamics are negligible within the exposure duration of a single frame. Alternatively, the dynamics in the scene may be faster than the frame rate of the camera, such that the interaction between changing scene radiance and measurements have to be explicitly modeled [34].

B. Calibration

The approaches described above model the functional relationship between sensor measurements and scene intensity. However, the actual parameters of the models need to be known, for example, the matrix \mathbf{H} in the general linear model, the kernel K in the convolutional model, or the row and column operators Φ_L and Φ_R in the separable model. When the mask design and sensor–mask distance are known precisely, these parameters may be known from the design process. However, typically, fabrication techniques are not sufficiently accurate, and one must measure the system response of the actual system as built; this process is called calibration.

We refer to the situation where the system parameters are known as model-based calibration. In this case, a computational model of the lensless imaging system design and appropriate wavefront (or light intensity) propagation models can be used to estimate/predict the calibration parameters computationally. The main advantage of the technique is that no actual measurements from the device are needed. The disadvantage is that any errors or differences between the design and the actual device are not accounted for, resulting in reconstruction artifacts.

When the system parameters are measured physically, we call it experimental calibration. This process typically involves taking one or more images of a “known” scene with the device and using the acquired measurements and the “known” scene to recover the calibration parameters. The simplest known scene to use is a point source, and hence the calibration measurements may be called PSF measurements. Depending on the forward model approximation (Section 6.A), the number of calibration measurements required

varies greatly, as described below. Hence, approximations to the forward model reduce not only computational requirements for reconstruction, but also significantly reduce the number of calibration measurements required.

Calibrating the general linear model. With a general linear model, one must measure a unique PSF for every point in the 2D or 3D scene. For large-pixel-count cameras, this means millions of calibration images are required. Given a 2D measurement for each PSF, this means the calibration procedure must estimate N^2 parameters (where N is the number of pixels). Hence, except for extremely low-resolution models, experimental calibration of general linear model parameters is intractable. As a consequence, the only viable approach for general linear models is model-based calibration, assuming that the device parameters are exactly known *a priori*—and perfectly realized.

Calibrating convolutional models. The ease of calibration is one of the primary advantages of the convolutional model. Since the PSF is shift invariant at each depth, only a single PSF capture is required for calibration. In the 3D case, depth-dependent scaling can also be modeled by a physical approximation; however, more accurate results come from multiple images acquired as the point source is moved axially. Similarly, if the PSF is not shift invariant across the FoV, a local convolutional model may apply, where a coarse grid of PSFs is acquired at different lateral positions across the FoV. The number of required calibration images in this case depends on how quickly the PSF changes across the FoV [25].

Calibrating separable models. In a separable model, calibration amounts to estimating the row and column operators Φ_L and Φ_R . Typically, this is done by projecting line images (rows and columns), rather than point sources, and capturing a series of lensless images with the line being translated to cover the entire FoV. This is slightly more challenging than calibration with a convolutional model—but since the number of images required remains limited, it is practically feasible, unlike the general linear model.

C. Algorithms for Reconstruction

Consider a particular lensless imaging system and assume that we have chosen an appropriate forward model to characterize the system, performed the calibration required, and estimated the calibration parameters. When a lensless image of an unknown scene is acquired, the measurements obtained \mathbf{y} are also known. Without loss of generality, assume that the forward model can be represented in functional form as $f_{\mathbf{H}}(\mathbf{x})$, where \mathbf{x} is the unknown scene parameters, and \mathbf{H} is the calibration parameters that have been estimated and are considered known. The goal of the reconstruction algorithm is to estimate the unknown image \mathbf{x} , given both the lensless measurements \mathbf{y} and the calibration parameters \mathbf{H} .

1. Convex Optimization

The image reconstruction problem can generically be cast as an optimization problem, as

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{y} - f_{\mathbf{H}}(\mathbf{x})\|_2 + \lambda \mathcal{R}(\mathbf{x}), \quad (7)$$

where λ is a Lagrange multiplier, and $\mathcal{R}(\mathbf{x})$ is a regularizer that typically imposes priors on the estimated image.

In most lensless imaging systems, the function $f_{\mathbf{H}}$ is linear, and with an appropriate choice of regularizer terms, the optimization

problem can be made convex. In such instances, a whole host of algorithms have been developed to solve these large-scale convex optimization problems [101,102]—and depending on the particulars of the choice of regularizer and forward model, one of these generic convex optimization algorithms is utilized.

The most common regularizers used are those that enforce smoothness [Tikhonov regularization, total variation (TV), transform domain sparsity] or those that enforce sparsity in the reconstruction (L_1 or L_0 norm). Typically, smoothness priors are more appropriate for natural scene reconstructions, while for many biological applications (especially those involving fluorescence), sparsity and related priors/regularizers are more appropriate.

Computational complexity and reconstruction run time.

The optimization problem above is a large-scale problem. Even for moderate resolution (say one-megapixel images), the optimization problem typically has $\approx 10^6$ unknowns to be estimated and $> 10^6$ equations. Consequently, direct implementations of traditional optimization techniques may be too slow to meet the rigorous run-time requirements in some applications. This poses strict constraints on the computational burden, especially in applications that require real-time performance (≈ 30 fps) or those that need the reconstruction algorithms to be implemented on-device. In almost all of these examples, the faster forward models such as convolutional or separable models are used; this choice, in conjunction with fast optimization algorithms (e.g., Wiener deconvolution) and their efficient (and if needed, parallel) implementations on hybrid edge compute platforms (including CPUs and GPUs), allows for these application-dependent needs to be met. Several of these lensless imaging devices have been demonstrated to achieve 30–100 fps real-time performance with an image resolution of a few megapixels, indicating that with the right choice of forward model and reconstruction algorithm, the computational complexity challenge can be reasonably handled.

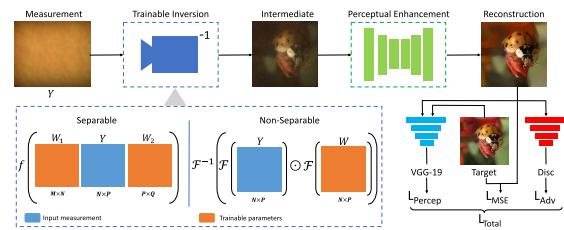
2. Data-Driven Techniques

Over the last decade, advances in machine learning, especially deep learning, have revolutionized inverse algorithms in many areas, and lensless imaging is no exception. Data-driven techniques provide a few significant advantages when compared to traditional optimization-based algorithms. First, these algorithms are less reliant on physical models of light transport and, therefore, better able to account for model errors. Second, these algorithms can be tuned to particular statistics of application domains, which allows performance improvements. Third, the direct utilization of these algorithms allows us to exploit the many radical advances in fast computing and associated hardware advances. Finally, these data-driven techniques rely on prior statistics about the reconstructed scene, resulting in significant performance improvements over optimization-alone techniques.

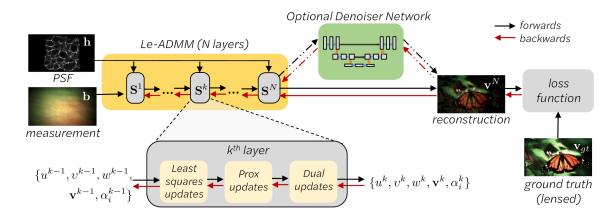
Data for learning. The data for learning can be generated by either experimentally capturing or simulating the measurements. Experimentally capturing training data [103–105] allows for incorporating all the non-idealities present in a real system into the data, thereby making the method generalize for various real scenarios. On the other hand, simulating the training data is much more efficient. When simulating measurements, performance improves when more aspects of the optics are incorporated (e.g., diffraction, sensor properties, spectral response) [106].

Deep learning architectures and algorithms. Various approaches perform two steps in the reconstruction network.

A With trainable inversion of lensless forward model



B With unrolled network of model-based optimization



C With deep back-projection network enhancement

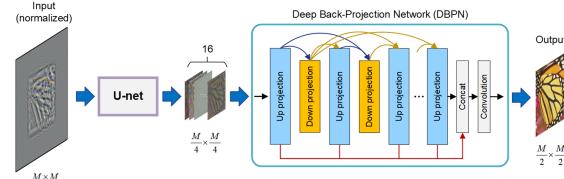


Fig. 8. Data-driven approaches to reconstruct lensless images: (A) modified from [103], (B) modified from [104], (C) modified from [106].

The first part incorporates the lensless system model to convert measurements to the image space, and the second part enhances the fidelity of the reconstruction. Monakhova *et al.* [104] used an unrolled network of alternating direction method of multipliers (Le-ADMM, Fig. 8B), where the iterative steps of ADMM are unrolled into layers, to simultaneously optimize the ADMM parameters and reconstruct the image (Fig. 9A). An optional U-net based denoiser network follows this first step. Khan *et al.* [103,105] used a first part that mimics the camera model formulation with trainable parameters. This first part is called the trainable inversion layer that maps the highly multiplexed input to the image space (Fig. 8A). The second part is U-net that perceptually enhances the final reconstruction to produce photo-realistic images (Fig. 9B). Wu *et al.* [106] used a deep backprojection network (DBPN) as the second part (Fig. 8C) to improve the reconstruction quality. The DBPN [107] has repeated up- and downsampling layers that provide an error feedback mechanism that realizes self-correcting of features and enhances the resolution of the output image.

7. APPLICATIONS

A. Photography, Refocusing, and 3D Imaging

In photography-based applications, it is not possible to control the placement of objects in the scene with respect to the sensor. Additionally, the objects are usually lit with ambient illumination, such as indoor lighting. In such cases, mask-modulated lensless systems encode the incoherent scene information onto the sensor, and the image is reconstructed using one of the algorithms mentioned in the previous section.

All-in-focus 2D imaging or photography was shown for objects more than a foot or 30 cm away from the lensless system [22–24, 42,77]. Particularly, DeWeert *et al.* [21] reconstructed images of a cityscape. For closer scene distance, refocusing was shown with a

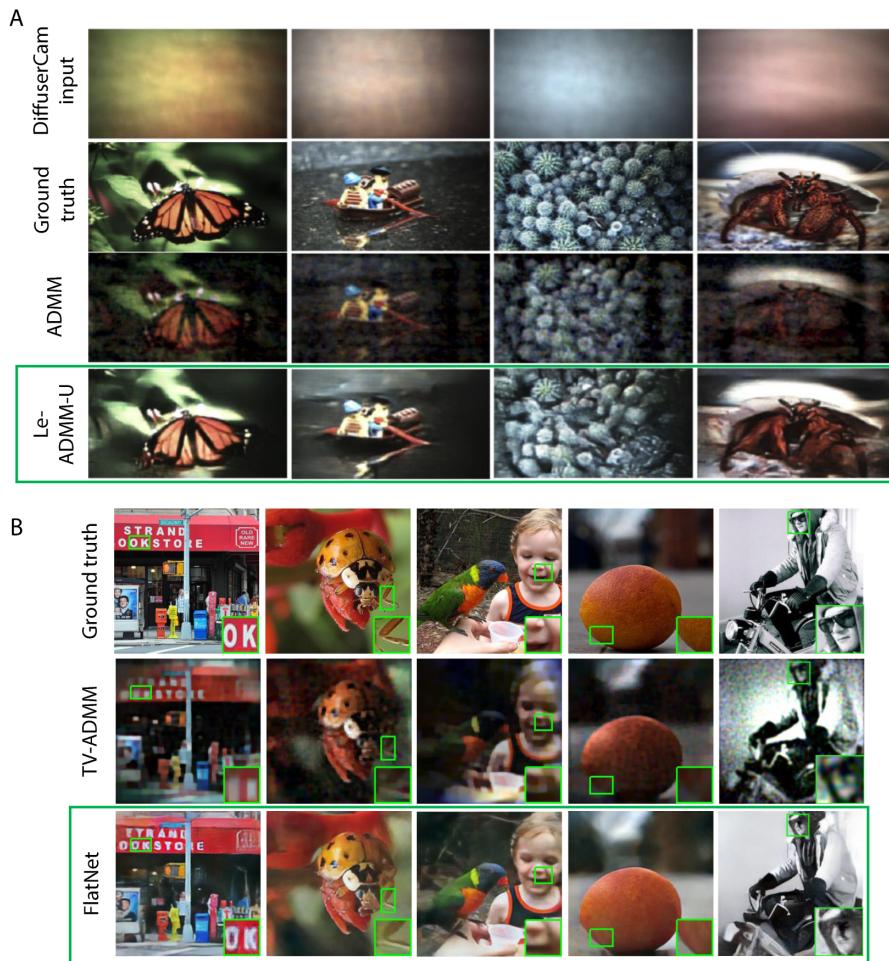


Fig. 9. Reconstruction results using data-driven techniques. ADMM and TV-ADMM use iterative convex optimization techniques. Le-ADMM-U [104] and FlatNet [103], outlined in green box, use feed-forward neural network, trained with data, to reconstruct. The data-driven methods drastically outperform, in quality, the optimization techniques. (A) Modified from [104] and (B) modified from [103].

fixed mask [23,77,108] and with a programmable mask [46]. For much closer scenes, 3D imaging was shown [23,53]. Exploiting the compressive nature of lensless imaging, the following applications were shown from just from a single capture/measurement: 3D imaging [23,53], video [34], and spectral imaging [33].

B. Microscopy

Microscopy is one area that benefits from the small form factor and unique design space provided by lensless imaging. In traditional microscopes, there is a practical trade-off among FoV, resolution, and light collection efficiency: large NA objective lenses that collect the most light and have the highest resolution are also larger and heavier than lower-NA objective lenses, as well as having a smaller FoV [27]. However, the trade-offs among FoV, resolution, and light collection efficiency for lensless imaging are quite different from those in a traditional microscope. In principle, the FoV of a lensless microscope can be expanded by creating a larger sensor or sensor array, without sacrificing resolution. Indeed, lensless microscopy has demonstrated large FoV high-resolution imaging in a number of contexts [19,27,62,109–112]. Additional advantages of lensless imaging for microscopy include 3D reconstruction abilities and reduced size, weight, and cost. These are important considerations for pre-clinical studies of freely moving animals

that cannot carry heavy microscopes and for clinical use where small size and weight can facilitate point-of-care use. However, one practical challenge for lensless microscopy is achieving high-quality chromatic filtering when operating in fluorescence mode. In traditional fluorescence microscopy, the excitation light can be filtered with a high extinction ratio using multilayer dielectric filters in the collimated light path. In lensless imaging, however, scattered excitation and fluorescence emission light are incident on the device from a wide range of angles because the sample is placed close to the device. As a result, multilayer filters do not perform well since their passband wavelengths are strongly dependent on the angle of incidence of incoming light. Despite these challenges, recent work has demonstrated that it is possible to reconstruct high-quality fluorescent images *in vivo* using a combination of dielectric and absorption filters [25,26]. While these demonstrations show that fluorescence lensless imaging is possible *in vivo*, more work is needed to achieve the same extinction ratios and corresponding SNRs as traditional lensed microscopes.

C. Emerging Applications

The size reduction achievable by a lensless configuration leads to many applications where space is severely constrained. One such application is endoscopy, where lensless technologies enable

imaging through an optical fiber bundle. Porat *et al.* [113] and Shekel *et al.* [114] exploited speckle-correlation properties (illumination modulated) to image through a fiber bundle. In another work, Shin *et al.* [115] placed an amplitude-coded aperture in front of the fiber bundle to capture mask-modulated measurements and reconstruct them. The current demonstrations of lensless endoscopy have been on a small scale and as a proof-of-concept and need further exploration to bring this technology to clinical use.

The heavily multiplexed measurements from mask-modulated lensless systems are visually unrecognizable, and reconstructing an image requires intimate knowledge of the system parameters. However, current imaging models of (non-holographic) lensless systems are simple, retain all the information, and are hackable to retrieve private information. Tan *et al.* [116] explored adding more analog operations on the sensor, such as pooling and quantization, as an attempt to destroy private information (e.g., face identity) while retaining useful information (e.g., face detection). On the other hand, software was used to remove sensitive information from post-capture measurement [117]. However, further exploration is needed to bring the privacy-enabled lensless system to practical use.

8. FUTURE OUTLOOK

Lensless imaging offers considerable advantages in size, weight, and form factor compared to traditional imaging systems. That makes lensless imaging an attractive option for a variety of space-constrained applications. Under-display cameras, micro-robotic vision, implantable sensors, and virtual/augmented reality headsets are some of the upcoming exciting applications that can benefit from the form factor of lensless cameras. In addition, an array of lensless cameras can be tiled with flexible interconnects to create a flexible imaging surface that can conform to various shapes. Such flexible imaging surfaces can benefit a range of applications such as health-monitoring sleeves, full-brain imaging, and 360°-view study of animals and plants, to name a few examples.

Computation is integral in extracting information from lensless measurements and adds significantly to the power requirements of the imaging system. Since many space-constrained applications are also power constrained, efficient computational implementation is an important area for improvement. The computational power problem could be alleviated in a few ways. First, lensless measurements can be wirelessly transmitted to the cloud, where powerful algorithms extract information. Second, efficient algorithms with a low-resource footprint can be developed. Third, a dedicated system-on-a-chip (SoC) can be developed to perform low-power lensless image processing. Additionally, jointly designing algorithms and SoCs can achieve incredible low-power performance.

Beyond imaging, lensless systems also have the potential for vision and inference applications. While recent works have focused on improving imaging performance, optimizing lensless systems for direct vision tasks is an exciting opportunity for future work. Face detection and verification, fingerprint identification, gesture recognition, human counting, and object tracking are some examples of vision applications that could immediately benefit from lensless systems.

Recent advances in lensless imaging have shown great promise in adding an exciting new dimension to the design of imaging systems. The lensless paradigm lays an additional pillar in technology that can enable new and novel future applications.

Funding. National Science Foundation (1730574, IIS-1652633, DMR-1548924); Defense Advanced Research Projects Agency (N66001-17-C-4012); National Institutes of Health (RF1NS110501); Weill Neurohub.

Disclosures. The authors declare no conflicts of interest.

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

REFERENCES

1. I. Al-Haytham, *Book of Optics (Kitab Al-Manazir)* (1011).
2. J. Al-Khalili, "In retrospect: Book of Optics," *Nature* **518**, 164–165 (2015).
3. J. Zahn, *Oculus artificialis teledioptricus sive telescopium (The Eye Also of a Telescope or an Artificial Teledioptricus)* (Johann Christoph Lochner, 1702).
4. M. Daniel, "Daguerre (1787–1851) and the Invention of Photography," (2004).
5. R. J. Easby, "Early photography: Niépce, Talbot and Muybridge," (2015).
6. A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: the KITTI dataset," *Int. J. Robot. Res.* **32**, 1231–1237 (2013).
7. M. Cornacchia, K. Ozcan, Y. Zheng, and S. Velipasalar, "A survey on activity detection and classification using wearable sensors," *IEEE Sens. J.* **17**, 386–403 (2016).
8. R. Eliakim, Z. Fireman, I. M. Gralnek, K. Yassin, M. Waterman, Y. Kopelman, J. Lachter, B. Koslowsky, and S. N. Adler, "Evaluation of the PillCam Colon capsule in the detection of colonic pathology: results of the first multicenter, prospective, comparative study," *Endoscopy* **38**, 963–970 (2006).
9. S.-E. Wei, J. Saragih, T. Simon, A. W. Harley, S. Lombardi, M. Perdoch, A. Hypes, D. Wang, H. Badino, and Y. Sheikh, "VR facial animation via multiview image translation," *ACM Trans. Graph.* **38**, 67 (2019).
10. Z. Ren, J. Meng, and J. Yuan, "Depth camera based hand gesture recognition and its applications in human-computer-interaction," in *8th International Conference on Information, Communications & Signal Processing* (IEEE, 2011), pp. 1–5.
11. J. Tan, L. Niu, J. K. Adams, V. Boominathan, J. T. Robinson, R. G. Baraniuk, and A. Veeraraghavan, "Face detection and verification using lensless cameras," *IEEE Trans. Comput. Imaging* **5**, 180–194 (2018).
12. V. Boominathan, J. K. Adams, M. S. Asif, B. W. Avants, J. T. Robinson, R. G. Baraniuk, A. C. Sankaranarayanan, and A. Veeraraghavan, "Lensless imaging: a computational renaissance," *IEEE Signal Process. Mag.* **33**(5), 23–35 (2016).
13. R. Hartley and S. B. Kang, "Parameter-free radial distortion correction with center of distortion estimation," *IEEE Trans. Pattern Anal. Mach. Intell.* **29**, 1309–1321 (2007).
14. P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *ACM SIGGRAPH 2008 classes on—SIGGRAPH* (ACM, 2008), p. 1.
15. N. Wadhwa, R. Garg, D. E. Jacobs, B. E. Feldman, N. Kanazawa, R. Carroll, Y. Movshovitz-Attias, J. T. Barron, Y. Pritch, and M. Levoy, "Synthetic depth-of-field with a single-camera mobile phone," *ACM Trans. Graph.* **37**, 1–13 (2018).
16. A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2005), Vol. **2**, pp. 60–65.
17. O. Liba, K. Murthy, Y.-T. Tsai, T. Brooks, T. Xue, N. Karnad, Q. He, J. T. Barron, D. Sharlet, R. Geiss, S. W. Hasinoff, Y. Pritch, and M. Levoy, "Handheld mobile photography in very low light," *ACM Trans. Graph.* **38**, 1–16 (2019).
18. S. Nayar, "Computational cameras: redefining the image," *Computer* **39**, 30–38 (2006).
19. A. Ozcan and U. Demirci, "Ultra wide-field lens-free monitoring of cells on-chip," *Lab Chip* **8**, 98–106 (2008).
20. A. Ozcan and E. McLeod, "Lensless imaging and sensing," *Annu. Rev. Biomed. Eng.* **18**, 77–102 (2016).
21. M. J. DeWeert and B. P. Farm, "Lensless coded-aperture imaging with separable Doubly-Toeplitz masks," *Opt. Eng.* **54**, 023102 (2015).
22. M. S. Asif, A. Ayremiou, A. Sankaranarayanan, A. Veeraraghavan, and R. G. Baraniuk, "FlatCam: thin, lensless cameras using coded aperture and computation," *IEEE Trans. Comput. Imaging* **3**, 384–397 (2017).

23. V. Boominathan, J. K. Adams, J. T. Robinson, and A. Veeraraghavan, "PhlatCam: designed phase-mask based thin lensless camera," *IEEE Trans. Pattern Anal. Mach. Intell.* **42**, 1618–1629 (2020).
24. G. Kuo, N. Antipa, R. Ng, and L. Waller, "DiffuserCam: diffuser-based lensless cameras," in *Imaging and Applied Optics 2017 (3D, A/O, COSI, IS, MATH, pcAOP)*, OSA Technical Digest (Optical Society of America, 2017), , paper CTu3B.2.
25. G. Kuo, F. L. Liu, I. Grossrubatscher, R. Ng, and L. Waller, "On-chip fluorescence microscopy with a random microlens diffuser," *Opt. Express* **28**, 8384–8399 (2020).
26. J. Adams, V. Boominathan, S. Gao, A. Rodriguez, D. Yan, C. Kemere, A. Veeraraghavan, and J. Robinson, "In vivo imaging of vasculature and neural activity with a flat, lensless microscope," *Nature Biomedical Engineering* (to be published).
27. J. K. Adams, V. Boominathan, B. W. Avants, D. G. Vercosa, F. Ye, R. G. Baraniuk, J. T. Robinson, and A. Veeraraghavan, "Single-frame 3D fluorescence microscopy with ultraminiature lensless FlatScope," *Sci. Adv.* **3**, e1701548 (2017).
28. I. Reshetouski, H. Oyaizu, K. Nakamura, R. Satoh, S. Ushiki, R. Tadano, A. Ito, and J. Murayama, "Lensless imaging with focusing sparse URA masks in long-wave infrared and its application for human detection," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (2020), Vol. **12364 LNCS**, pp. 237–253.
29. E. Candès, "Compressive sampling," in *International Congress of Mathematicians, Madrid, August 22–30, 2006*, (European Mathematical Society, 2006), Vol. **3**, pp. 1433–1452.
30. E. Candès and M. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.* **25**(2), 21–30 (2008).
31. K. Yanny, N. Antipa, W. Liberti, S. Dehaeck, K. Monakhova, F. L. Liu, K. Shen, R. Ng, and L. Waller, "Miniscope3D: optimized single-shot miniature 3D fluorescence microscopy," *Light Sci. Appl.* **9**, 171 (2020).
32. F. Linda Liu, G. Kuo, N. Antipa, K. Yanny, and L. Waller, "Fourier DiffuserScope: single-shot 3D Fourier light field microscopy with a diffuser," *Opt. Express* **28**, 28969–28986 (2020).
33. K. Monakhova, K. Yanny, N. Aggarwal, and L. Waller, "Spectral DiffuserCam: lensless snapshot hyperspectral imaging with a spectral filter array," *Optica* **7**, 1298–1307 (2020).
34. N. Antipa, P. Oare, E. Bostan, R. Ng, and L. Waller, "Video from stills: lensless imaging with rolling shutter," in *IEEE International Conference on Computational Photography (ICCP)* (Institute of Electrical and Electronics Engineers, 2019).
35. J. Richardson, L. Grant, and R. Henderson, "Low dark count single-photon avalanche diode structure compatible with standard nanometer scale CMOS technology," *IEEE Photon. Technol. Lett.* **21**, 1020–1022 (2009).
36. A. Kirmani, H. Jeelani, V. Montazerhodjat, and V. K. Goyal, "Diffuse imaging: creating optical images with unfocused time-resolved illumination and sensing," *IEEE Signal Process. Lett.* **19**, 31–34 (2012).
37. D. Wu, G. Wetzstein, C. Barsi, T. Willwacher, Q. Dai, and R. Raskar, "Ultra-fast lensless computational imaging through 5D frequency analysis of time-resolved light transport," *Int. J. Comput. Vis.* **110**, 128–140 (2014).
38. G. Satat, M. Tancik, and R. Raskar, "Lensless imaging with compressive ultrafast sensing," *IEEE Trans. Comput. Imaging* **3**, 398–407 (2017).
39. G. Kim, K. Isaacson, R. Palmer, and R. Menon, "Lensless photography with only an image sensor," *Appl. Opt.* **56**, 6450–6456 (2017).
40. E. Caroli, J. B. Stephen, G. Di Cocco, L. Natalucci, and A. Spizzichino, "Coded aperture imaging in X- and gamma-ray astronomy," *Space Sci. Rev.* **45**, 349–403 (1987).
41. K. Tajima, T. Shimano, Y. Nakamura, M. Sao, and T. Hoshizawa, "Lensless light-field imaging with multi-phased Fresnel zone aperture," in *IEEE International Conference on Computational Photography (ICCP)* (IEEE, 2017), pp. 1–7.
42. J. Wu, H. Zhang, W. Zhang, G. Jin, L. Cao, and G. Barbastathis, "Single-shot lensless imaging with Fresnel zone aperture and incoherent illumination," *Light Sci. Appl.* **9**, 53 (2020).
43. D.-K. Yang and S.-T. Wu, *Fundamentals of Liquid Crystal Devices* (Wiley, 2014).
44. G. Huang, H. Jiang, K. Matthews, and P. Wilford, "Lensless imaging by compressive sensing," in *IEEE International Conference on Image Processing (ICIP)* (2013), pp. 2101–2105.
45. A. Zomet and S. K. Nayar, "Lensless imaging with a controllable aperture," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2006), Vol. **1**, pp. 339–346.
46. Y. Hua, S. Nakamura, M. S. Asif, and A. C. Sankaranarayanan, "SweepCam—depth-aware lensless imaging using programmable masks," *IEEE Trans. Pattern Anal. Mach. Intell.* **42**, 1606–1617 (2020).
47. Y. Wu, M. K. Sharma, and A. Veeraraghavan, "WISH: wavefront imaging sensor with high resolution," *Light Sci. Appl.* **8**, 2047–7538 (2019).
48. M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Process. Mag.* **25**(2), 83–91 (2008).
49. X. Yuan, G. Huang, H. Jiang, and P. A. Wilford, "Block-wise lensless compressive camera," in *IEEE International Conference on Image Processing (ICIP)* (IEEE, 2017), pp. 31–35.
50. S. Moazeni, E. H. Pollmann, V. Boominathan, F. A. Cardoso, J. T. Robinson, A. Veeraraghavan, and K. L. Shepard, "A mechanically flexible implantable neural interface for computational imaging and optogenetic stimulation over 5.4×5.4 mm 2FoV," in *Digest of Technical Papers—IEEE International Solid-State Circuits Conference* (2021), Vol. **64**, pp. 288–290.
51. Y. Wu and A. Ozcan, "Lensless digital holographic microscopy and its applications in biomedicine and environmental monitoring," *Methods* **136**, 4–16 (2018).
52. P. R. Gill and D. G. Stork, "Lensless ultra-miniature imagers using odd-symmetry spiral phase gratings," in *Optics InfoBase Conference Papers* (Optical Society of America, 2013), paper CW4C.3.
53. N. Antipa, G. Kuo, R. Heckel, B. Mildenhall, E. Bostan, R. Ng, and L. Waller, "DiffuserCam: lensless single-exposure 3D imaging," *Optica* **5**, 1–9 (2018).
54. D. Lange, C. W. Storment, C. A. Conley, and G. T. Kovacs, "A microfluidic shadow imaging system for the study of the nematode *Caenorhabditis elegans* in space," *Sens. Actuators B* **107**, 904–914 (2005).
55. S. Seo, T. W. Su, D. K. Tseng, A. Erlinger, and A. Ozcan, "Lensfree holographic imaging for on-chip cytometry and diagnostics," *Lab Chip* **9**, 777–787 (2009).
56. X. Heng, D. Erickson, L. R. Baugh, Z. Yaqoob, P. W. Sternberg, D. Psaltis, and C. Yang, "Optofluidic microscopy—a method for implementing a high resolution optical microscope on a chip," *Lab Chip* **6**, 1274–1276 (2006).
57. X. Cui, L. M. Lee, X. Heng, W. Zhong, P. W. Sternberg, D. Psaltis, and C. Yang, "Lensless high-resolution on-chip optofluidic microscopes for *Caenorhabditis elegans* and cell imaging," *Proc. Natl. Acad. Sci. USA* **105**, 10670–10675 (2008).
58. G. Zheng, S. A. Lee, S. Yang, and C. Yang, "Sub-pixel resolving optofluidic microscope for on-chip cell imaging," *Lab Chip* **10**, 3125–3129 (2010).
59. D. Gabor, "A new microscopic principle," *Nature* **161**, 777–778 (1948).
60. Z. Gorocs and A. Ozcan, "On-chip biomedical imaging," *IEEE Rev. Biomed. Eng.* **6**, 29–46 (2013).
61. J. W. Goodman, *Statistical Optics*, 2nd ed. (Wiley, 2015).
62. O. Mudanyali, D. Tseng, C. Oh, S. O. Isikman, I. Sençan, W. Bishara, C. Oztoprak, S. Seo, B. Khademhosseini, and A. Ozcan, "Compact, lightweight and cost-effective microscope based on lensless incoherent holography for telemedicine applications," *Lab Chip* **10**, 1417–1428 (2010).
63. M. K. Kim, "Principles and techniques of digital holographic microscopy," *SPIE Rev.* **1**, 1–51 (2010).
64. U. Schnars, C. Falldorf, J. Watson, and W. Jüptner, *Digital Holography* (Springer, 2015), pp. 39–68.
65. J. W. Goodman, *Introduction to Fourier Optics*, 3rd ed. (Roberts & Co, 2004).
66. W. Bishara, U. Sikora, O. Mudanyali, T.-W. Su, O. Yaglidere, S. Luckhart, and A. Ozcan, "Holographic pixel super-resolution in portable lensless on-chip microscopy using a fiber-optic array," *Lab Chip* **11**, 1276–1279 (2011).
67. A. Greenbaum, Y. Zhang, A. Feizi, P.-L. Chung, W. Luo, S. R. Kandukuri, and A. Ozcan, "Wide-field computational imaging of pathology slides using lens-free on-chip microscopy," *Sci. Transl. Med.* **6**, 267ra175 (2014).
68. W. Luo, A. Greenbaum, Y. Zhang, and A. Ozcan, "Synthetic aperture-based on-chip microscopy," *Light Sci. Appl.* **4**, e261 (2015).

69. A. Greenbaum, U. Sikora, and A. Ozcan, "Field-portable wide-field microscopy of dense samples using multi-height pixel super-resolution based lensfree imaging," *Lab Chip* **12**, 1242 (2012).
70. A. Greenbaum and A. Ozcan, "Maskless imaging of dense samples using pixel super-resolution based multi-height lensfree on-chip microscopy," *Opt. Express* **20**, 3129–3143 (2012).
71. J. R. Fienup, "Phase retrieval algorithms: a comparison," *Appl. Opt.* **21**, 2758–2769 (1982).
72. R. Horraud, M. Hansard, G. Evangelidis, and C. Ménier, "An overview of depth cameras and range scanners based on time-of-flight technologies," *Mach. Vis. Appl.* **27**, 1005–1020 (2016).
73. S. Cova, A. Longoni, and A. Andreoni, "Towards picosecond resolution with single-photon avalanche diodes," *Rev. Sci. Instrum.* **52**, 408–412 (1981).
74. A. Velten, T. Willwacher, O. Gupta, A. Veeraraghavan, M. G. Bawendi, and R. Raskar, "Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging," *Nat. Commun.* **3**, 745 (2012).
75. M. O'Toole, D. B. Lindell, and G. Wetzstein, "Confocal non-line-of-sight imaging based on the light-cone transform," *Nature* **555**, 338–341 (2018).
76. "Hamamatsu photonics: streak camera," <https://www.hamamatsu.com/us/en/product/photometry-systems/streak-camera/operating-principle/index.html>.
77. T. Shimano, Y. Nakamura, K. Tajima, M. Sao, and T. Hoshizawa, "Lensless light-field imaging with Fresnel zone aperture: quasi-coherent coding," *Appl. Opt.* **57**, 2841–2850 (2018).
78. D. G. Stork and P. R. Gill, "Lensless ultra-miniature CMOS computational imagers and sensors," in *7th International Conference on Sensor Technologies and Applications* (2013), pp. 186–190.
79. D. G. Stork and P. R. Gill, "Optical, mathematical, and computational foundations of lensless ultra-miniature diffractive imagers and sensors," *Int. J. Adv. Syst. Meas.* **7**, 201–208 (2014).
80. A. K. Singh, G. Pedrini, M. Takeda, and W. Osten, "Scatter-plate microscope for lensless microscopy with diffraction limited resolution," *Sci. Rep.* **7**, 10687 (2017).
81. R. H. Dicke, "Scatter-hole cameras for x-rays and gamma rays," *Astrophys. J.* **153**, L101 (1968).
82. Y. Nakamura, T. Shimano, K. Tajima, M. Sao, and T. Hoshizawa, "Lensless light-field imaging with Fresnel zone aperture," Technical report (2016).
83. D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory* **52**, 1289–1306 (2006).
84. I. Reshetouski, R. Tadano, H. Oyaizu, K. Nakamura, and J. Murayama, "Lensless mismatched aspect ratio imaging," in *IEEE International Conference on Computational Photography (ICCP)* (IEEE, 2021).
85. M. Born, E. Wolf, A. B. Bhatia, P. C. Clemmow, D. Gabor, A. R. Stokes, A. M. Taylor, P. A. Wayman, and W. L. Wilcock, *Principles of Optics* (Cambridge University, 1999).
86. "Nanoscribe gmbh," <https://www.nanoscribe.de/>.
87. P. R. Gill, J. Tringali, A. Schneider, S. Kabir, D. G. Stork, E. Erickson, and M. Kellam, "Thermal Escher sensors: pixel-efficient lensless imagers based on tiled optics," in *Optics InfoBase Conference Papers* (OSA, 2017), Vol. **Part F46-C**, paper CTu3B.3.
88. F. Tian, J. Hu, and W. Yang, "GEOMScope: large field-of-view 3D lensless microscopy with low computational complexity," *Laser Photon. Rev.* **15**, 2100072 (2021).
89. W. Chi and N. George, "Phase-coded aperture for optical imaging," *Opt. Commun.* **282**, 2110–2117 (2009).
90. W. Chi and N. George, "Optical imaging with phase-coded aperture," *Opt. Express* **19**, 4294–4300 (2011).
91. J. R. Fienup, "Lensless coherent imaging by phase retrieval with an illumination pattern constraint," *Opt. Express* **14**, 498–508 (2006).
92. E. E. Fenimore and T. M. Cannon, "Coded aperture imaging with uniformly redundant arrays," *Appl. Opt.* **17**, 337–347 (1978).
93. S. R. Gottesman and E. E. Fenimore, "New family of binary arrays for coded aperture imaging," *Appl. Opt.* **28**, 4344–4352 (1989).
94. S. W. Golomb, *Shift Register Sequences* (Aegean Park, 1967).
95. A.-J. Fresnel, "Calcul de l'intensité de la Lumière au centre de l'ombre d'un Écran et d'une Ouverture circulaires éclairées par un point radieux," in *Mémoires de l'Académie des sciences de l'Institut de France*, (Imprimerie royale, 1821), chap. Note I., pp. 456–464.
96. H. Jiang, G. Huang, and P. Wilford, "Multi-view in lensless compressive imaging," in *Picture Coding Symposium (PCS)* (2014).
97. X. Yuan and Y. Pu, "Parallel lensless compressive imaging via deep convolutional neural networks," *Opt. Express* **26**, 1962–1977 (2018).
98. J. R. Miller, C. Y. Wang, C. D. Keating, and Z. Liu, "Particle-based reconfigurable scattering masks for lensless imaging," *ACS Nano* **14**, 13038–13046 (2020).
99. M. Levoy and P. Hanrahan, *Light Field Rendering* (ACM, 1996), pp. 31–42.
100. S. J. Gortler, R. Grzeszczuk, R. Szeliski, and M. F. Cohen, *The Lumigraph* (ACM, 1996), pp. 43–54.
101. A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imaging Sci.* **2**, 183–202 (2009).
102. B. Stephen, P. Neal, C. Eric, P. Borja, and E. Jonathan, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.* **3**, 1–122 (2011).
103. S. S. Khan, V. Sundar, V. Boominathan, A. Veeraraghavan, and K. Mitra, "FlatNet: towards photorealistic scene reconstruction from lensless measurements," *IEEE Trans. Pattern Anal. Mach. Intell.* (2020).
104. K. Monakhova, J. Yurtsever, G. Kuo, N. Antipa, K. Yanny, and L. Waller, "Learned reconstructions for practical mask-based lensless imaging," *Opt. Express* **27**, 28075–28090 (2019).
105. S. S. Khan, R. V. Adarsh, V. Boominathan, J. Tan, A. Veeraraghavan, and K. Mitra, "Towards photorealistic reconstruction of highly multiplexed lensless images," *IEEE International Conference on Computer Vision*, October, 2019, pp. 7859–7868.
106. J. Wu, L. Cao, and G. Barbastathis, "DNN-FZA camera: a deep learning approach toward broadband FZA lensless imaging," *Opt. Lett.* **46**, 130–133 (2021).
107. M. Haris, G. Shakhnarovich, and N. Ukita, "Deep back-projection networks for super-resolution," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2018), pp. 1664–1673.
108. J. Tan, V. Boominathan, A. Veeraraghavan, and R. Baraniuk, "Flat focus: depth of field analysis for the FlatCam lensless imaging system," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE, 2017), pp. 6473–6477.
109. W. Bishara, T.-W. Su, A. F. Coskun, and A. Ozcan, "Lensfree on-chip microscopy over a wide field-of-view using pixel super-resolution," *Opt. Express* **18**, 11181–11191 (2010).
110. S. Jiang, J. Zhu, P. Song, C. Guo, Z. Bian, R. Wang, Y. Huang, S. Wang, H. Zhang, and G. Zheng, "Wide-field, high-resolution lensless on-chip microscopy: via near-field blind ptychographic modulation," *Lab Chip* **20**, 1058–1065 (2020).
111. M. Sanz, J. Á. Picazo-Bueno, L. Granero, J. Garcíá, and V. Micó, "Compact, cost-effective and field-portable microscope prototype based on MISHELF microscopy," *Sci. Rep.* **7**, 43291 (2017).
112. H. Tobon-Maya, S. Zapata-Valencia, E. Zora-Guzmán, C. Buitrago-Duque, and J. García-Sucerquia, "Open-source, cost-effective, portable, 3D-printed digital lensless holographic microscope," *Appl. Opt.* **60**, A205–A214 (2021).
113. A. Porat, E. R. Andresen, H. Rigneault, D. Oron, S. Gigan, and O. Katz, "Widefield lensless imaging through a fiber bundle via speckle correlations," *Opt. Express* **24**, 16835–16855 (2016).
114. N. Shekel and O. Katz, "Using fiber-bending-generated speckles for improved working distance and background rejection in lensless micro-endoscopy," *Opt. Lett.* **45**, 4288–4291 (2020).
115. J. Shin, D. N. Tran, J. R. Stroud, S. Chin, T. D. Tran, and M. A. Foster, "A minimally invasive lens-free computational microendoscope," *Sci. Adv.* **5**, eaaw5595 (2019).
116. J. Tan, S. S. Khan, V. Boominathan, J. Byrne, R. Baraniuk, K. Mitra, and A. Veeraraghavan, "CANOPIC: pre-digital privacy-enhancing encodings for computer vision," *IEEE International Conference on Multimedia and Exposition*, July, 2020.
117. T. Nguyen Canh and H. Nagahara, "Deep compressive sensing for visual privacy protection in FlatCam imaging," in *International Conference on Computer Vision Workshop (ICCVW)* (Institute of Electrical and Electronics Engineers, 2019), pp. 3978–3986.