Lensless Imaging and Sensing

Aydogan Ozcan\textsuperscript{1,2,3} and Euan McLeod\textsuperscript{4}

\textsuperscript{1}Department of Electrical Engineering, \textsuperscript{2}Department of Bioengineering, and \textsuperscript{3}California NanoSystems Institute, University of California, Los Angeles, California 90095; email: ozcan@ucla.edu

\textsuperscript{4}College of Optical Sciences, University of Arizona, Tucson, Arizona 85721; email: euanmc@optics.arizona.edu

\textbf{Abstract}

High-resolution optical microscopy has traditionally relied on high-magnification and high-numerical aperture objective lenses. In contrast, lensless microscopy can provide high-resolution images without the use of any focusing lenses, offering the advantages of a large field of view, high resolution, cost-effectiveness, portability, and depth-resolved three-dimensional (3D) imaging. Here we review various approaches to lensless imaging, as well as its applications in biosensing, diagnostics, and cytometry. These approaches include shadow imaging, fluorescence, holography, superresolution 3D imaging, iterative phase recovery, and color imaging. These approaches share a reliance on computational techniques, which are typically necessary to reconstruct meaningful images from the raw data captured by digital image sensors. When these approaches are combined with physical innovations in sample preparation and fabrication, lensless imaging can be used to image and sense cells, viruses, nanoparticles, and biomolecules. We conclude by discussing several ways in which lensless imaging and sensing might develop in the near future.

\textbf{Keywords}

microscopy, lab-on-chip, holography, phase recovery, 3D imaging
1. INTRODUCTION

Historically, bioimaging at the microscopic scale has been performed using lenses, usually the compound microscope and its adaptations (1). Recently, however, imaging without lenses has matured as a modality competitive with traditional lens-based microscopy. In lensless microscopy, a diffraction pattern resulting from an object (based on, e.g., scattering or fluorescence) is recorded directly on a digital image sensor array without being optically imaged or magnified by any lens elements (Figure 1a). This recorded diffraction pattern is then computationally reconstructed to form an “image” of the object. The recent maturation of lensless imaging was made possible largely by the mass production of inexpensive digital image sensors with small pixel size and high pixel counts (2, 3), along with improvements in computing power and reconstruction algorithms used to process the captured diffraction patterns. Here we review recent progress on lensless imaging and sensing in domains including bright-field imaging, cytometry, holography, phase recovery, and fluorescence, as well as the quantitative sensing of specific sample properties derived from such images.

Compared with conventional lens-based microscopy, lensless approaches impart several key advantages: a large space–bandwidth product (large field of view and high resolution simultaneously; see, e.g., Figure 1b), cost-effectiveness, portability (Figure 1c,d), and depth-resolved three-dimensional (3D) imaging. These advantages make lensless imaging particularly well suited to analysis applications requiring large statistics. These applications include cytometry “needle-in-a-haystack” diagnostic tasks such as the Papanicolaou smear test for cervical cancer (Figure 1b) (4–7) and blood smear inspection for malaria diagnosis (8, 9), in which the high-resolution inspection of a large number of cells is required because parasitemia is typically less than 1%. Thus, many cells need to be screened to declare a sample “negative” with any reasonable level
of confidence. More generally, lensless imaging can be a good choice for any type of study that demands statistically significant estimates of a population on the basis of the analysis of a limited sample. Examples include the examination of thousands of sperm trajectories to identify rare types of motion (10, 11) and the performance of a complete blood count (12, 13). Tests designed for samples that have large ranges in number density (concentration) also benefit from imaging approaches with a large space–bandwidth product and depth of field, as it is possible to detect either a few objects or many thousands to millions of objects in a single image. Air quality and water quality tests are two examples of tests in which number density can vary widely.

Other categories of applications that are particularly well suited to lensless imaging are point of care and global health. In an effort to speed up diagnosis and reduce its cost via minimizing reliance on central laboratory testing and diagnostic facilities, these applications demand devices and approaches that are compact, portable, robust, cost-effective, and realistic for widespread distribution. Lensless imaging is a good option for point-of-care and global health applications because it does not require expensive precision microscope objective lenses and light sources. In many instances, individual light-emitting diodes (LEDs) are sufficient for illumination, and the most expensive component in the system is the image sensor, whose cost can be as low as a few tens of dollars due to the mass production of complementary metal-oxide semiconductor (CMOS) image sensors for use in mobile phones (2, 3). Lensless imaging is also compatible with cell phone and smartphone platforms (Figure 1d), which are generally suitable for point-of-care- and global health–related imaging, cytometry, and sensing applications (14–35).
2. LENSLESS IMAGING APPROACHES

2.1. Shadow Imaging

The simplest form of lensless imaging is shadow imaging, which requires only the basic hardware setup (Figure 1a) without the need for any image reconstruction procedures. In shadow imaging, a transparent sample is illuminated with a spatially limited light source from a distance of \( z_1 \), and the shadow of the sample is recorded directly on the image sensor at a distance of \( z_2 \). Due to diffraction over the finite distance \( z_2 \) and the relatively large pixel size at the sensor array, the recorded shadows are pixelated, are not in focus, and may consist of either blurry spots when the light source is incoherent (Figure 2a, c, d) (36) or concentric fringe interference patterns when the light source is coherent (Figure 2b, e) (37). The coherence of the light source depends on both the lateral size of the light source relative to \( z_1 \) (spatial coherence) and the bandwidth of the spectrum of the light source (temporal coherence) (38). In Section 2.3, we explain how the coherence of the light source can be exploited to reconstruct “images” of the sample. In this section, we rely only on the shadow information, without an attempt to reconstruct images from these undersampled diffraction patterns on the sensor array.

Because typical biological objects of interest are partially transparent, the recorded shadows are not black and white but rather exhibit grayscale complexity. This complexity can be used to identify specific types of particles based on a so-called pattern-matching approach. This approach was initially used to perform blood cell counts to differentiate among red blood cells, fibroblasts, murine embryonic stem cells, hepatocytes, and polystyrene beads of similar sizes (36, 39). Many other types of cells have also since been imaged using lensless shadow imaging; these include leukocytes, HeLa cells, sperm cells, A549 cells, RWPE-1 prostate epithelial cells, Madin–Darby canine kidney cells, cardiomyocytes, human alveolar epithelial cells, human mesenchymal stem cells, Pseudomonas aeruginosa, Schizosaccharomyces pombe, NIH 3T3 cells, MCF-7 cells, bioluminescent Escherichia coli, and HepG2 cells (40–53). The lensless imaging approach can monitor cell division, motility, and viability, among other properties (47, 50, 54), and can operate within standard cell culture incubators (Figure 3).

For cytometry applications in general, the physical resemblance between the cell morphology and its shadow is not needed, because specific recognition of the cell type is the primary goal.
However, to obtain shadows that bear the closest possible resemblance to the original objects, it is desirable to minimize $z_2$, thereby minimizing the effect of diffraction (e.g., compare panels c and d of Figure 2). Typically, the smallest practical value of $z_2$ is determined by the layer of glass placed on the sensor during manufacturing in order to protect it from dust and scratches. The thickness of this protective glass ranges from 10 $\mu$m to several hundred micrometers. Even using sensors with sample-to-sensor distances of up to 4 mm, usable shadow information can be obtained for cytometric analysis.

Even if $z_2$ is minimized by selecting an appropriate sensor-array design, the bottleneck for the resolution of a shadow imaging–based technique will be dictated by the pixel size, which varies between $\sim$1 $\mu$m and $\sim$10 $\mu$m for various CMOS and charge-coupled device (CCD) chips. This undersampling limitation can be improved by placing an array of small apertures directly below the sample. In this configuration, diffraction between the sample and the apertures is minimized, and the intensity of the light transmitted through each aperture can be monitored separately. When the sample is “scanned” relative to the aperture array, assuming that it is a thin sample, the transmittance through each aperture can approximately correspond to the localized transmittance of a single part of the sample at that instant in time. In this way, the aperture array functions similarly to an array of scanning probes, as in the case of, for instance, near-field scanning optical microscopy (NSOM). The resolution of this approach for two-dimensional (2D) objects is limited by the size of the apertures and by the precision with which the sample can be scanned relative to the apertures (which is more typically the practical limiting factor). One way in which the scanning can be performed is through flowing the object(s) of interest along a microfluidic channel fabricated on top of the aperture array, thus forming an “optofluidic” microscope. Such a shadow imaging approach has proven to provide submicrometer resolution of *Caenorhabditis elegans*, *Giardia lamblia*, and other microscale objects (55, 56).

### 2.2. Fluorescence Imaging

Lensless fluorescence imaging (57, 58) is conceptually similar to shadow imaging, except that (a) the scattering-based partially coherent or coherent signal is now replaced with incoherent fluorescent emission from the sample and (b) the excitation light and captured emission light are of different wavelengths. The experimental setup thus requires an emission filter to be placed...
between the sample and the sensor to reject the excitation wavelength [unless the sensor is not sensitive to the excitation wavelength (59)]. For many fluorescence applications, even with an emission filter, signal-to-background levels are still a concern, and extra measures can be helpful in reducing leakage of excitation light. These measures can include excitation in a total internal reflection geometry (57, 58, 60) or fabrication of specialized CMOS sensor arrays with integrated filters (61–63), where fluorophore densities as small as 100 molecules/µm² have been detected.

As with other imaging approaches, the resolution of lensless fluorescence imaging is conventionally limited by the point spread function (PSF) of the system, which is the size of the spot recorded on the image sensor that comes from a single fluorescent point/emitter in the sample plane. Note that this PSF is a strong function of sample height, and one of the limiting factors is diffraction over \( z_2 \). As an example, for a typical \( z_2 \) value of \( \sim 200 \) µm, the resolution that is dictated by the PSF is \( \sim 200 \) µm (57, 59). This resolution is significantly worse than that of most lens-based microscope systems. Thus, although there are clear advantages in terms of cost, portability, and field of view for a fluorescent lensless imaging setup, the space–bandwidth product obtained using this approach does not initially present a major advantage over that of conventional imaging.

However, it is possible to partially overcome this resolution limitation using hardware and/or computational approaches. One hardware approach, as mentioned above, involves fabricating the filter elements directly into the individual pixels of the sensor array such that \( z_2 \) can be minimized, given that no additional filter element is necessary between the sample and the sensor planes. With this approach, a spatial resolution on the order of \( \sim 13 \) µm has been achieved (62). Another hardware approach involves using a tapered faceplate (a tightly packed array of optical fibers) to relay the fluorescence emission (60). In this configuration, a tightly spaced bundle of fibers is placed in contact with the sample. As these fibers relay the light toward the image sensor, the fiber bundle enlarges, providing moderate lensless on-chip magnification. The emission filter can then be placed between the fiber taper and the image sensor. With this setup, the resolution can be as small as 4 µm, determined by the pitch of the fiber-optic taper where it contacts the sample. Yet another hardware approach involves placing a nanostructured mask in close proximity to the object of interest (64, 65). This nanostructured mask generates a PSF that depends on the location of the object/emitter. In other words, instead of the PSF of the system being position independent (i.e., spatially invariant), its shape now depends sensitively on the pattern of nanostructuring that is in close proximity to the fluorescent object plane, allowing the precise locations of nearby objects to be computed from the structure of the recorded emission spot. With this approach, subpixel-level resolutions as good as 2–4 µm have been achieved using compressive decoding (64, 65).

Finally, structured excitation can also be used to improve resolution. If the excitation spot is tightly localized and its position is known with high accuracy, then the resolution of the system is limited by the precision with which the excitation spot can be scanned. One way that this technique can be implemented is through the Talbot effect, which has been used to enable fluorescence imaging with a resolution of 1.2 µm on a chip (66). Structured illumination can be also be employed to image and perform cytometry through moderately scattering media, such as large volumes of whole blood (13).

Computational methods can be used to further improve the apparent resolution of the image after it has been captured. One of the most common ways to deblur an image is via deconvolution—for example, with the Lucy–Richardson algorithm (57, 67, 68). With this approach, resolutions as good as 40–50 µm have been achieved (57). Even better resolution can be obtained through compressive decoding, which can yield a resolution of \( \sim 10 \) µm (58). This approach relies on the natural sparsity of fluorescent objects and has been used to successfully image fluorescent transgenic C. elegans worms (63) in a dual-mode lensfree microscope that merges fluorescence and bright-field modes. A compressive decoding approach has also been used to demultiplex
fluorescence emission channels in samples labeled with multiple colors (69), as well as in the nanostructured on-chip imaging approach discussed in the previous paragraph (65).

### 2.3. Digital Holographic Reconstruction in Lensless On-Chip Microscopy

As mentioned in Section 2.1, when the aperture of the light source is sufficiently small and monochromatic, the optical field impinging on the sample plane can be considered coherent, and the recorded shadows exhibit interference fringe patterns (Figure 2k,e). From these interference fringes, one can infer the optical phase of the scattered light, making it possible to reconstruct the optical field that is transmitted through the objects on the sample and thereby providing an in-focus “image” of the sample (Figure 4).

More precisely, the interference fringe pattern recorded on the sensor is an in-line hologram (70, 71), which is the intensity pattern generated by the interference between light scattered from an object on the sample and a reference wave that that passes undisturbed through the transparent substrate. At the plane where this light interacts with the sample, we can describe its electric field as the sum of the reference wave at the sample, $E_{R,s}$, and the object wave at the sample, $E_{O,s}$:

$$E_i = E_{R,s} + E_{O,s} = A_R + A_O(x, y)e^{i\phi_O(x, y)},$$

(1)

where $A_R$ is the amplitude of the spatially uniform plane reference wave (because $z_1 \gg z_2$, the incident wave can be considered a plane wave over the range between the sample and sensor), $A_O(x, y)$ is the spatially varying amplitude (transmittance) of the object, and $\phi_O(x, y)$ is the spatially varying phase (optical thickness) of the object, which is assumed to be a thin sample for the sake of simplicity. The goal of digital holographic reconstruction is to recover $A_O$ and $\phi_O$ on the basis of measurement(s) of the light intensity further downstream at the image sensor plane.

From these downstream measurements, it is possible to reconstruct the amplitude and phase of the object by using the angular spectrum approach (72). This approach consists of computing the Fourier transform of the captured hologram, multiplying it by the transfer function of free space, and then inverse Fourier transforming. Mathematically,

$$E_o = \mathcal{F}^{-1}[\mathcal{F}(E_i(x, y)) \times H_{\lambda}(x, y)],$$

(2)

where $E_o$ is the reconstructed optical field of the object, $E_i(x, y)$ is the captured hologram, and $H_{\lambda}(x, y)$ is the transfer function of free space ($n = 1$):

$$H_{\lambda}(x, y) = \begin{cases} e^{ikz_2\left(\frac{1}{2f_2^2} - \frac{1}{2f_o^2}\right)}, & f_2^2 + f_o^2 < \frac{1}{\lambda^2} \\ 0, & f_2^2 + f_o^2 \geq \frac{1}{\lambda^2} \end{cases}$$

(3)

Here, $\lambda$ is the wavelength of the light, $k = 2\pi/\lambda$, $f_i$ and $f_o$ are spatial frequencies, and $z_2$ is the same sample-to-sensor distance shown in Figure 1a. Often $z_2$ may not be precisely known before the capture of the hologram, and some computational “refocusing” can be performed to estimate it and provide the sharpest reconstructions.

In contrast to the shadow imaging presented in Section 2.1, two of the significant advantages of digital holographic reconstruction are the improvements in resolution and detection signal-to-noise ratio (SNR) (73). These improvements are visible in Figure 4, where grating objects cannot be directly resolved from the raw shadows but are visible upon holographic reconstruction. This resolution improvement helps not only in discerning two nearby objects, such as lines on a grating, but also in imaging finer features on a single object (see, e.g., Figure 4f,g).

There are also some limitations to the method of this basic implementation of digital holographic reconstruction, although many of these limitations may be overcome or mitigated through
Digital holographic lensless imaging. (a) Raw hologram image of a grating with 1-µm half-pitch. (b) Pixel-superresolved hologram of the same grating. (Inset) Holographic fringes corresponding to high spatial frequencies become visible in the superresolved hologram. (c) Reconstruction of the amplitude transmittance from the low-resolution hologram in panel a. The grating lines are not resolved. (d) Reconstruction of the amplitude transmittance from the superresolved hologram in panel b. The grating lines are now resolved. (e) Reconstruction of the amplitude transmittance (orange) of a different grating with 225-nm half-pitch, demonstrating the ability of lensless imaging to resolve extremely fine features. (f) Lensless amplitude image of red blood cells, one of which is infected with a malaria parasite (*Plasmodium falciparum*). (g) Lensless phase image corresponding to panel f, where the malaria parasite appears more distinctly. (h) Conventional 40× objective microscope image, where the infected cell is indicated with an arrow. Panels a–d modified from Reference 80. Panel e modified from Reference 82. Panels f–h modified from Reference 8.

Alternative approaches presented throughout the rest of Section 2. One limitation is that the resolution is limited by the pixel size of the sensor, which is ~1–2 µm for current mass-produced sensors. Methods to overcome this limitation are discussed in Section 2.4. A second limitation is the presence of the twin-image artifact that is a consequence of only being able to measure the intensity of the field at the image sensor plane. Overcoming this limitation is discussed in Section 2.6. Third, it is necessary that the scattered wave be weak compared with the reference wave, which restricts this simple form of holographic reconstruction to relatively sparse samples. However, this limitation has also been overcome, as discussed further in Section 2.6. Fourth, the requirement of coherence between the illumination wave and the scattered wave means that it is...
not immediately possible to combine this approach with fluorescence, unless some changes are
made to the experimental setup. Finally, color imaging is not trivial due to the temporal coherence
requirement that the light source be monochromatic; this topic is addressed in Section 2.7 with
several approaches that can provide color information in holographically reconstructed lensfree
images.

2.4. From Pixel-Limited Resolution to Signal-to-Noise
Ratio–Limited Resolution

In a conventional microscope, the resolution is typically limited by the numerical aperture (NA)
of the objective lens following the expression \( \Delta x \sim \lambda/(2NA) \). Without immersion media, it is
not possible to utilize lenses with \( NA > 1 \). This limiting value of NA results in the oft-quoted
conventional diffraction limit of \( \Delta x > \lambda/2 \).

In holographic lensless on-chip imaging, the resolution of the computationally reconstructed
image is equal to the resolution of the captured hologram. Because there is no lens with finite NA
to limit the resolution, the resolution-limiting factor is, in theory, the refractive index \( n \) of the
medium that fills in the space between the sample and sensor planes and, in practice, the pixel
size of the sensor. The drive to miniaturize image sensors for use in mobile phones has helped
reduce the pixel size on manufactured sensors; however, the smallest commercially available pixel
sizes are still larger than 1 \( \mu \)m. This native pixel size would result in resolutions comparable to
conventional microscope objectives, with \( NA \sim 0.25 \) for 500-nm light. Although this resolution
can be sufficient for many applications, there has recently been a strong push to improve the
resolution of lensless imaging beyond this limit so as to extend the range of applications and
capabilities of on-chip microscopy tools.

Assuming a sensor could have arbitrarily small pixels, there would still be a diffraction-imposed
limit on the resolution of the captured hologram. Because \( z_2 \) is larger than several wavelengths,
disturbances in the electric field with spatial frequency components greater than \( 1/\lambda \) resulting
from fine spatial features in the object will evanescently decay, whereas only those disturbances
with characteristic size greater than approximately \( \lambda/2 \) will propagate to the far field (as shown
mathematically in Equation 3). Note, however, that although these spatial frequencies decay, they
are not lost completely. Furthermore, from the knowledge that the electric field must physically
be described as a continuous function, it is theoretically possible, at least for a space-limited object,
to recover these “lost” spatial frequencies through analytic continuation (74, 75). In practice, the
success of these approaches depends on the measurement’s SNR and object support information
and, therefore, have been limited in their use.

Additionally, the temporal and spatial coherence properties of the light source can indepen-
dently limit the resolution of lensless holographic imaging. The coherence of the system deter-
mines the maximum angle at which light scattered from an object will produce an observable
interference pattern at the detector (38). This maximum scattering angle corresponds to the NA
of this lensless imaging system and its half-pitch resolution:

\[
\Delta x = \frac{\lambda}{2NA} = \frac{\lambda}{2\pi \sin \theta_{\text{max}}}.
\]

(4)

The maximum scattering angle \( (\theta_{\text{max}}) \) may be limited by the temporal and/or spatial coherence
of the system. The temporal coherence is determined primarily by the spectral bandwidth of the
system \( \lambda \), which implies the following coherence length (38):

\[
L_{\text{coh}} = \left( \frac{2 \ln 2}{\pi} \right) \frac{\lambda^2}{n\Delta \lambda}.
\]

(5)
A coherent interference pattern is generated when the phase delay between a reference ray and a scattered ray is no longer than the coherence length, implying

$$\theta_{\text{max}} \leq \arccos \left( \frac{z_2}{z_2 + L_{\text{coh}}} \right).$$  \hspace{1cm} (6)

Substituting this expression into Equation 4 yields the limit to the resolution imposed by temporal coherence of illumination.

The spatial coherence is influenced primarily by the lateral size of the light source $D$ and the source-to-sample distance $z_1$, assuming $z_1 \gg z_2$, as is the case for an on-chip imaging design. The resulting maximum scattering angle is given by

$$\theta_{\text{max}} \leq \arctan \left( \frac{0.61 \lambda z_1 / D}{z_2} \right),$$  \hspace{1cm} (7)

where the numerator of the arctangent is one-half the distance to the first zero of the complex coherence factor of the light source, observed at the sample plane (38). When substituted into Equation 4, this value provides the spatial coherence imposed limit on resolution.

Taken together, the above equations can be used to prescribe the necessary properties of the light source to achieve a desired resolution. For example, if we ignore the pixel size–induced undersampling and resolution limitation (addressed in the following paragraphs) to achieve a resolution equivalent to that from a 0.8-NA microscope objective using 500-nm light in air with $z_2 = 100$ µm, the spectral bandwidth of the source must be $\Delta \lambda < 1.6$ nm, and the ratio $D/z_1$ must be $<2.2 \times 10^{-3}$. These levels of coherence do not require a laser source and can be attained through the combination of an LED, a narrowband interference filter, and a large-diameter pinhole or multimode optical fiber with $z_1$ on the order of a few centimeters. Exceedingly coherent sources, such as lasers, can ultimately degrade image quality through speckle artifacts or spurious interference fringes resulting from multiple reflections within and by the substrate (76).

Even with a coherent or partially coherent light source, there still remains a “resolution gap” between the sensor pixel size of $\sim 1$ µm and the “diffraction limit” of $\lambda/2$. One particularly successful approach to bridge this gap is pixel superresolution. In this technique, multiple captured images of the same object that are shifted with respect to each other in the $x$ and $y$ directions by noninteger numbers of pixels can be synthesized into a single high-resolution image. This technique enables many images with large pixel sizes to be converted into an image with much smaller virtual pixels. Note that this approach is not simply interpolation, or upsampling, which can never truly improve resolution; rather, it allows one to recover fine features that are normally undersampled. The first step of this procedure involves estimating the shifts between the images with subpixel accuracy. This step can be done using a variety of “optical flow” techniques, such as iterative gradient-based techniques (77, 78). Next, the raw images are upsampled to the desired new high resolution; the intermediate values are initially determined through interpolation of neighboring low-resolution pixels. One way of computing the pixel-superresolved image, termed the shift-and-add method, is as follows (79):

$$\hat{z} = \left( \sum_{k=1}^{N} F_k^T D^T D F_k \right)^{-1} \left( \sum_{k=1}^{N} F_k^T D^T y_k \right).$$  \hspace{1cm} (8)

Here, the $F_k$ are block-circulant matrices that describe the relative shifts between the images, $D$ is the decimation (downsampling) operator that converts a high-resolution image to a low-resolution image, and the $y_k$ are the interpolated upsampled captured images, represented (reshaped) as a column vector. The high-resolution estimate $\hat{z}$ is likewise represented as a column vector. In this equation, the second term is conceptually the sum of the shifted raw images, whereas the first term
is a normalization parameter based on the number of shifted images, their degree of redundancy, and the desired amount of resolution enhancement. In addition to this shift-and-add method of pixel superresolution, other methods exist, such as iterative gradient-descent optimization routines (77).

Whereas these equations are written with shifts of the image sensor in mind, it is often more practical to shift the light source rather than the object (80), although a flowing object within a microfluidic channel can also be used (81). This is because larger and less-precise shifts of the light source can be used to generate subpixel shifts on the image sensor due to the large $z_1/z_2$ ratio. This approach can also be implemented in field-portable devices where, instead of a single light source being shifted, multiple, spatially separated light sources are sequentially activated (8).

Pixel superresolution has been successfully combined with lensless holographic on-chip microscopy to obtain half-pitch resolutions as fine as 225 nm on commercially available CMOS sensors with a pixel size of 1.12 µm (Figure 4e) (82, 83). In this approach, a high-resolution hologram (e.g., that in Figure 4b) is synthesized using pixel superresolution before the image is computationally reconstructed. Because the resolution of the reconstruction equals the resolution of the hologram, the capture of fine spatial features in the hologram allows fine features on the object to be recovered. With this approach, lensless holographic imaging has been made nearly diffraction limited, with half-pitch resolutions equivalent to those from 0.8–0.9-NA microscope objectives (82, 84).

Although pixel superresolution has enabled high-resolution imaging in combination with small-pixel CMOS sensors, its value is even more apparent when combined with large-pixel CCD sensors. CCD sensors are typically lower noise and can have larger pixel counts than CMOS sensors; however, their pixel sizes are often greater than ~4–5 µm, which limits their resolution when used as a lensless imaging device (Figure 2c,d). Through the use of pixel superresolution in combination with CCD sensors, half-pitch resolutions as fine as 0.84 µm have been achieved (82). This technique generates gigapixel images with more than one billion pixels that carry useful (i.e., resolvable) information. In other words, the smallest resolvable feature and the characteristic linear dimension of the field of view differ by more than four orders of magnitude.

The performance of pixel superresolution can be further improved if the spatial responsivity of each pixel is known at a microscopic scale. This is known as the pixel function, which is not uniform because of the presence of, for instance, wiring and microscopic circuit elements present at each pixel and the potential for a microlens to be fabricated on top of each pixel. Several approaches can be used to determine the pixel function of a sensor a posteriori. One method is to scan a tightly focused laser beam across a pixel and record its response. This method is reasonable for larger-size pixels such as those on CCD sensors (82). However, for ~1-µm pixels, the diffraction-limited spot size of the laser beam is approximately equal to that of the pixel size, and useful information cannot be obtained unless a near-field scanning probe is utilized. In cases where experimental measurements of the pixel function become difficult, one can employ alternative computational approaches based on blind deconvolution routines or optimization routines (82). Here, a known object with fine features is holographically imaged and reconstructed with an initial guess for the pixel function (e.g., a guess that the pixel function is uniform across the pixel). Then the guess for the pixel function is modified iteratively. Modifications that lead to an improvement in the reconstruction quality are pursued until convergence is achieved. As an example, this approach was used in the reconstruction of Figure 4e.

In addition to the holographic lensless imaging approach, subpixel resolving techniques have also been used in other lensless imaging approaches. In lensless microscopes based on near-contact-mode imaging, pixel superresolution techniques have been used to achieve 0.75-µm resolution of objects carried by the microfluidic flow (85). This approach can be combined with machine
learning to discriminate between different types of cells (86). A similar contact imaging approach, termed ePetri, has recently been used to image static samples in near contact with the image sensor. Here, a shifting light source creates shadows that shift over fixed pixel locations (87, 88). The ePetri platform has been applied to the imaging of viral plaques (89) and waterborne parasites (90).

2.5. Three-Dimensional Lensless Imaging

Another advantage of the holographic lensless imaging approach is the ease with which images at different depths can be computationally generated through variation of the value of $z_2$ in Equation 2. This can be considered digital refocusing of a captured image. With this approach, it is possible to obtain some degree of 3D sectioning. Here, the depth of focus of a single object scales with $O^2/\lambda$, where $O$ is the object diameter. Practically, unless other 3D imaging techniques are utilized, the depth of focus results in an axial resolution of at least $\sim 40 \mu m$ (91–93). The typical range of depths of digital refocusing is on the order of millimeters. It is important to realize that with this simple approach, the sample must remain relatively sparse so that a clean reference wave is still available for holographic reconstruction.

Improved localization in $z$ of sparse objects is possible through triangulation using at least two light sources: one that illuminates the sample at normal incidence and another that illuminates the sample at oblique incidence (94). With this approach, localization accuracies of $\sim 300–400$ nm have been achieved (94). Additionally, for unique computational reconstructions of the recorded holograms, it is helpful if two different wavelengths (e.g., blue and red) are used for the two illuminations so that it is unambiguous which hologram comes from which light source. This technique has been successfully used to track and finely resolve the 3D trajectories of thousands of sperms in a volume of $\sim 5.0 \times 3.5 \times 0.5$ mm in a single experiment (10, 11). This large number of statistics has been instrumental in identifying rare and unique swimming patterns in human and animal sperm, such as the finding that of the human sperm that travel in helical paths in vitro, $\sim 90\%$ travel in a right-handed helix as opposed to a left-handed helix.

When many light sources are used sequentially from many different angles, lensless tomodographic microscopy can also be performed. Here, many holograms are recorded from significantly different angles. A single tomogram can be reconstructed from these multiple holograms via, for instance, filtered back-projection (91, 95). This approach has also been implemented in a compact portable platform (96), used to image C. elegans worms in three dimensions without the use of lenses (91, 97), and combined with a microfluidic device for imaging flowing samples, demonstrating the first implementation of an optofluidic tomographic microscope (98).


In the preceding sections, the coherent interference pattern captured on the sensor was regarded as an inline hologram—that is, the interference between a reference beam and a scattered object beam. This assumption enables one to infer the optical phase of the light wave (although this phase also included a twin-image artifact), which is at the core of the computational reconstruction of an image of the object. However, this coherent interference pattern can more generally be considered a coherent diffraction pattern. Under this more general consideration, where a clean reference wave is not necessarily present, the optical phase can no longer be inferred from a single measurement. However, with additional information—for example, either from a priori knowledge of the nature or size of the object (73) or from multiple measurements at different $z_2$ distances (4, 5, 99)—it
becomes possible to recover the lost phase information, which eliminates the source of the twin-image artifact and enables the high-fidelity reconstruction of dense connected samples, such as tissue samples used in clinical examination in pathology labs.

For samples that are still relatively sparse, the optical phase (without the twin-image artifact) can be recovered through the assumption that outside the boundaries of individual objects, the amplitude and phase of the light should be uniform. This additional boundary information regarding the sample is enough to enable the elimination of the twin-image artifact. One way to remove this artifact is to computationally propagate the optical field back and forth between the sensor (hologram) plane and the sample plane, enforcing known information at each plane (73). At the sensor plane, the known information is the captured amplitude of the optical field, whereas at the sample plane, the known information is the optical field outside the rough boundaries of the objects. After each propagation, components of the optical field corresponding to known information are replaced with that information, while the remaining components continue to be dynamically updated with each propagation. Experiments have shown that after several iterations of this procedure, the optical field converges to a physically consistent function, and it is possible to reconstruct the objects with minimal twin-image artifacts. A key parameter in the success of this approach is the ability to accurately define the boundaries of objects in the image (known as the object support). In practice, it can sometimes be a challenge to accurately define these regions in an automated fashion. Lens-based reflective imaging of the same sample has also been used to improve the selection of the object support in a lensless on-chip microscope (100).

The iterative phase recovery algorithm described above is only one example of the whole class of phase recovery algorithms, dating back to the Gerchberg–Saxton algorithm, which was pioneered in the 1970s (101, 102). In general, these algorithms involve computationally propagating an optical field between planes and enforcing any known information at each plane. These approaches can involve more than two planes, and the known information can take many different forms. A specific type of phase recovery approach that has shown widespread success is based on multiple amplitude or intensity measurements at various values of $z_2$ (4, 5, 99). This technique can be used to image dense, connected samples such as tissue slices (99) or blood vessel formation within engineered tissue models (103), which would not normally be recoverable using on-chip holography alone. Here, iterative propagation among the various planes is performed, enforcing the measured amplitude of the optical field at each plane. Once convergence is achieved, the optical field of any given measurement plane can be propagated to the sample plane to generate amplitude and phase images of the object. Because this approach does not rely on the presence of a clean reference wave, it works well even for dense, connected samples without the need for any object support information (Figure 5).

In these iterative phase recovery algorithms, several iterations are often required to achieve a desired level of convergence, which can take considerable computational time. One way to speed up convergence and computational time is through the transport-of-intensity equation (104, 105):

$$\frac{\partial I(x, y)}{\partial z} = -\frac{\lambda}{2\pi} \nabla \cdot [I(x, y)\nabla \phi(x, y)],$$

where $I(x, y)$ is the intensity of the optical field, $\nabla$ is the gradient operator in the $(x, y)$ plane, and $\phi$ is the optical phase. This equation serves as an alternative to, for instance, the coherent Rayleigh–Sommerfeld diffraction integral, which is often used to calculate the propagation of an optical field. Using the measured intensities at different planes as boundary conditions, one can numerically solve this partial differential equation to provide an estimate for the optical phase. With the transport-of-intensity equation, the computational time to convergence can be sped up considerably (99), and the quality of the reconstructed image remains high.
Figure 5
Tissue imaging using multilength phase recovery. (a) Full field-of-view (FOV) amplitude image of a stained human breast carcinoma tissue slice, reconstructed from measurements of diffracted intensity at different sample-to-sensor distances. (b) Magnified regions of interest from panel a clearly showing individual cells and their nuclei. (c) Comparison image of the same regions using a conventional bright-field microscope. (d) Examples of some of the raw diffracted intensities captured by the sensor. Abbreviation: NA, numerical aperture. Modified from Reference 99.

Another holographic on-chip microscopy approach to imaging of dense samples is lensfree imaging using synthetic aperture (LISA) (106). The experimental setup for LISA is similar to that used in lensless tomography; however, the samples in LISA are assumed to be 2D (as in a histopathology sample), and the processing algorithm is entirely different. By illuminating the sample at high angles of incidence, one can convert high spatial frequencies at the object to lower spatial frequencies (and vice versa) at the hologram (image sensor) plane. High-resolution reconstructions with an effective NA of 1.4 can be obtained by cycling through a sufficient number of incidence angles and orientations (together with source-shifting- or sample-shifting-based pixel superresolution), thus sampling a large region in the frequency domain. Furthermore, if there is some overlap (redundancy) between the images in the frequency domain, these separate measurements can be used to perform iterative phase recovery, enabling the imaging of dense samples (Figure 6).

2.7. Color Imaging in Lensless On-Chip Microscopy
The high-resolution imaging techniques described in Sections 2.3–2.6 all rely on the capture of coherent interference patterns on the sensor. When the illumination is incoherent, as discussed in Sections 2.1 and 2.2, the resolution is relatively poor. To obtain coherence in illumination, the light must be monochromatic. This requirement results in monochrome images, which can be less desirable from the point of view of, for instance, microscopists, pathologists, and medical practitioners. In many cases, valuable information is encoded in color from specific stains applied.
Figure 6

Tissue imaging using a lensfree synthetic aperture approach. The top left panel shows the captured raw hologram from a stained human breast cancer tissue for the full field of view (FOV) of the lensless imaging system. The other panels show magnified regions after reconstruction and colorization. Abbreviations: CMOS, complementary metal-oxide semiconductor; ROI, region of interest. Modified from Reference 106.

Figure 7

10 μm
10 μm
10 μm
10 μm
10 μm
10 μm
10 μm

Figure 7

Lensless color imaging methods. Each panel shows an image of two cells in a stained Papanicolaou smear. (a) The direct superposition of reconstructions from multiple color channels results in washed-out colors and rainbow-like artifacts. (b) Averaging image colors in the YUV color space can mitigate the rainbow artifacts. (c) Colorizing the image using Dijkstra’s shortest path method also mitigates rainbow artifacts; however, slight color variations within a single object are lost. (d) A comparison image acquired using a conventional 20× microscope objective is used as the gold standard for color imaging. Modified from Reference 6.

the sample. Another way to overcome these rainbow artifacts is Dijkstra’s shortest path algorithm, which is a graph-search algorithm aiming to find the shortest path from a given node to the remaining nodes. This approach has also been used for lensfree color imaging, yielding very similar results to those of the YUV color-space averaging approach discussed above (Figure 7c) (6, 108).

These color imaging approaches, among others, have been implemented in a variety of portable devices. Both YUV averaging and Dijkstra’s shortest path method have been successfully incorporated into a portable pixel superresolution–based on-chip microscopy device using different-colored LEDs (7). In shadow imaging, color imaging has been used to evaluate hemoglobin concentrations (109) and enzyme-linked immunosorbent assays (110). In the optofluidic microscope approach, in both a format where the object flows over small apertures and a format where the object flows without apertures but in close proximity to the image sensor (see Sections 2.1 and 2.4), color imaging has been performed, again using multiple LEDs (41, 111, 112). Color imaging has also been incorporated with the lensfree tomography approaches discussed in Section 2.5 (113).

3. Lensless Sensing

The lensfree computational imaging devices described above can also be converted into sensors when combined with specialized substrates or specialized photonic elements and automated image processing routines. In this section, we describe some examples of lensless sensing, including the fabrication and imaging of self-assembled nanolenses for the sensing of nanoscale particles and viruses, measurements of cellular mechanics, the use of metallic beads as labels for biochemical specificity, and sensing based on plasmonics.

3.1. Self-Assembled Nanolenses for Detection of Nanoparticles and Viruses over Large Fields of View

An ever-present challenge in biosensing and imaging is the ability to detect smaller and smaller objects (114, 115). The lensless imaging approaches described in Section 2 are well able to detect and image microscale objects such as single cells; however, the detection of nanoscale objects such as single viruses or nanoparticles is difficult. The primary reason is that the intensity of the electromagnetic wave scattered by such small particles is so weak compared with the background (reference) wave intensity that it can become lost in the background noise of these on-chip imaging systems. Indeed, individual spherical particles smaller than ~250 nm cannot be reliably discerned
from background noise via the above-discussed approaches, unless a significant refractive index contrast exists between the particle and the surrounding medium.

In order to reduce this detection limit to permit the sensing of virus-scale particles, a self-assembly approach has been used to fabricate liquid polymer nanoscale lenses around the target particles. These self-assembled nanolenses provide an increased scattering signal compared with the particles alone, partially due to the increased volume and partially due to their refractive geometry preferentially directing the scattered signal toward the image sensor. Three distinct approaches have been pursued to form these nanolenses, namely tilting-based formation (83, 116, 117), formation by solvent evaporation (118, 119), and formation by the condensation of a polymer vapor (120, 121).

In the tilting-based approach, the target nanoparticles are suspended in a Tris-PEG-HCl buffer. A drop of this suspension is placed on a plasma-treated glass surface and left to rest for a few minutes, after which the surface is gently tilted to let the bulk of the drop slide to the edge of the glass. In the wake of the drop, small droplets (nanolenses) of liquid are pinned around target particles that adhere to the glass. In this liquid, polyethylene glycol (PEG) is the critical component. PEG is a nontoxic water-soluble polymer, and at the molecular weight used here (600 Da), it evaporates very slowly in ambient conditions. This last property allows nanometer-scale droplets of PEG to remain stable for many minutes to hours, permitting the acquisition of multiple images. With this approach, nanoparticles and viruses with diameters as small as 100 nm can be sensed over large fields of view (e.g., 20–30 mm²) by use of lensless on-chip imaging (Figure 8a–d) (117).

Another approach to nanolens formation involves letting the initial drop evaporate instead of tilting it to the side. In this approach, the particles are again suspended in a liquid mixture. This mixture contains PEG, surfactant, and a volatile solvent. The volatile solvent evaporates relatively rapidly, leaving behind a nanofilm consisting mostly of PEG and surfactant. Due to surface tension, this nanofilm rises in the form of a meniscus around any particles that are sitting on the substrate and embedded in the film (118). The lensless sensing detection limit is again ~100 nm (119), similar to the tilting-based approach in terms of both limit of detection and field of view. However, in this approach, no particles are lost, as may happen when the excess liquid is removed via tilting. The physical nature of the nanolenses is also different in this approach in that the lenses are part of a continuous film coating, as opposed to isolated droplets.

The final approach to nanolens formation involves the condensation of a nanofilm from a vapor. One advantage of this technique is that particles can be initially deposited on the substrate via multiple means; they do not need to be first suspended in the same liquid that will form the nanolenses. This makes the condensation approach more compatible with particles having various forms of surface chemistry. After particle deposition, a continuous nanofilm of pure PEG is condensed on the substrate, and nanolenses form around the particles due to surface tension (Figure 8e). Another advantage of this technique is that it enables direct control over the thickness of the film that is formed. This control allows the size of the nanolens to be tuned to achieve the maximum possible signal, which has enabled the limit of detection to be reduced to ~40 nm (Figure 8f) (120). Furthermore, the optimum signal for each particle correlates strongly with the particle’s size. By using image processing routines to extract the optimum signal for each particle, an automated sensor platform has been able to measure particle-size distributions within the sample (Figure 8g, h). As such, the platform can be considered an alternative to dynamic light scattering (122, 123), nanoparticle tracking analysis (124), or electron microscopy for nanoparticle sizing. Additionally, this entire approach has been integrated into a compact, cost-effective, field-portable platform (weighing <500 g), taking advantage of the benefits of the lensfree on-chip imaging approach (Figure 8g).
3.2. Cell Capture, Manipulation, and Sensing

Many applications of lensless imaging involve the detection and monitoring of cells. Specialized substrates that enable the capture or manipulation of cells can aid in this task and provide extra functionality. One example is the use of a perforated transparent substrate that can be used to trap, for instance, HeLa cells through the use of an applied negative pressure. The number of trapped cells can then be counted using a lensless imaging setup (125). Similarly, small microwells can be used to trap individual cells (126). For selectivity, microarrays of antibodies can be used to capture specific cells for quantification (40). In another example, the active capture surface is integrated directly into the fabrication of the image sensor (51). Finally, a photoconductive substrate material can be used in an optoelectronic tweezer setup; when combined with lensless imaging, this approach enables real-time manipulation of a large number of cells across the entire active area of the image sensor (>20 mm²) (127).

3.3. Bead-Based Labeling for Specific Sensing and Cytometric Analysis in a Lensless Design

Specificity is extremely important for any type of biochemical sensor. Often, fluorescent molecules are used as labels that can be bound to specific biomolecules to show their presence and location. However, fluorescent labels are unsuitable for a high-resolution lensless holographic imaging setup.
because their emission is incoherent with their excitation. Instead, beads with biomolecularly functionalized surfaces can be used to provide specificity, as the scattering from these beads is coherent with their illumination. Metallic beads work especially well due to their strong scattering properties, which can be enhanced even further by taking advantage of their plasmonic resonances.

Gold and silver nanobeads have been used as specific labels to identify and count the number of CD4 and CD8 cells in a cell suspension (128). CD4 and CD8 cells are specific types of T lymphocytes, and their relative populations are important for evaluating the stage of human immunodeficiency virus (HIV) infection or AIDS, as well as for evaluating the efficacy of antiretroviral treatment. Counting the relative populations of these cells can be challenging, however, because the only significant difference between these cells is in the types of proteins expressed on their membranes. Under a microscope, both types of cells look virtually identical. In order to sense and count these cells in a lensless imaging setup, the authors of this study (128) used gold nanoparticles functionalized with anti-CD4 antibodies and silver nanoparticles functionalized with anti-CD8 as labels. After incubation of the cells with the nanoparticles, the CD4 cells became coated with gold nanoparticles and the CD8 cells became coated with silver nanoparticles. By comparing the spectral response of different cells under lensless imaging, the investigators were able to discriminate these two types of cells with greater than 95% accuracy using a machine learning algorithm.

Beads have also been used as labels for sensing DNA in a sandwich assay that is read out using lensless imaging (129). In this technique (Figure 9), a substrate is first functionalized with short strands of DNA that are complementary to the bases near one end of the target DNA sequence. Magnetic beads are also functionalized with short strands of DNA that are complementary to the bases at the other end of the target DNA sequence. First, a sample to be analyzed is incubated with the substrate; then, the substrate is incubated with the functionalized magnetic beads. If the target strands were present in the sample, they will be labeled with magnetic beads. These magnetic beads are then quantified using a lensfree imaging system. Another sandwich-type assay using beads has also been read out using a lensless system, although in this case, bead and fluid motion for the incubation and washing steps was controlled by surface acoustic wave transducers (130).

### Figure 9

DNA sensing using magnetic beads as labels. This sandwich assay enables the detection of individual strands of DNA in a lensless imaging system. (a) The substrate is functionalized with DNA complementary to part of the target DNA. (b) Target DNA binds to the functionalized surface. (c) Magnetic beads functionalized with complementary DNA bind to the free portion of the captured target DNA strand. (d) Unbound beads are washed away, and the remaining beads indicate the presence of individual target strands. Modified from Reference 129.

### 3.4. Lensless Plasmonic Sensing

Lensless imaging can be combined with plasmonically active substrates for sensing purposes (131). In one application, transmission through a plasmonic nanohole array was monitored using a
portable sensing apparatus (132, 133). Arrays of nanoscale holes in thin metal films had previously been observed to exhibit extraordinary optical transmission (134, 135); that is, the transmission through the array exceeded the predicted transmission through a single hole times the number of holes. This phenomenon occurs because the transmission is enhanced via surface plasmon waves propagating along the film. This extraordinary optical transmission exhibits a strong resonance peak at a specific wavelength. The resonance is quite sensitive to the refractive index of the material within the evanescent field of the nanohole array. If the nanohole array is biochemically functionalized to capture specific biomolecules, it can then act as a sensor through monitoring of the shift of the resonance peak. If the nanohole array is initially illuminated by a wavelength tuned to the resonance responding to no bound biomolecules, then when biomolecules start to bind to the sensor, a drop in transmission occurs, and this drop can be observed with a lensless imaging system (131, 132). In such a device, many plasmonic arrays functionalized for different analytes can be used in parallel to develop a massively multiplexed sensor. For improved sensitivity, two wavelengths can be used simultaneously: one tuned to the initial resonance and one tuned slightly in the direction the resonance will shift. The transmission ratio between the two wavelengths can be used to measure the degree of binding of the analyte with a sensitivity approximately a factor of two higher than that of the single-LED device (133). These and similar systems have been used to sense biomolecules (132, 133).

4. FUTURE OUTLOOK

Although lensless imaging has already demonstrated significant results in terms of its ability to capture large-field-of-view images with high resolution and image quality in a compact and cost-effective platform, there are still a number of areas for future improvement. One such area is that of resolution. Investigators have obtained a resolution equivalent to that from 1.4-NA microscope objectives (see Section 2.6), which is approximately as high as can be achieved using standard objective lenses. Yet this need not pose a hard limit for lensless imaging because, due to the computational imaging approach, resolution is more closely tied to SNR than to any of the physical parameters of the system. Through higher-SNR systems and new algorithms capable of extrapolating high–spatial frequency information from low-frequency knowledge of the complete field (e.g., through analytic continuation), it should be possible to further enhance resolution, with the assumption that a high index medium (e.g., an oil) fills in the space between the sample and sensor planes.

Complementary to resolution is sensitivity. Current lensless on-chip imaging approaches have a sensitivity limit of ~250 nm without nanolenses, and ~40 nm using self-assembled nanolenses. With future improvements to imaging SNR, as well as potential improvements to sample preparation and nanolens formation, it should be possible to reduce both of these limits.

Another area of improvement is that of imaging speed. Approaches such as the pixel super-resolution discussed in Section 2.4 and the multihight phase recovery discussed in Section 2.6 require the capture of many images in sequence, which can limit their application in systems with rapid dynamics. The use of high-speed image sensors with bright light sources can solve this deficiency. In addition to the speed of image acquisition, the duration of image processing and computational reconstruction steps can sometimes pose a challenge in certain applications. These speeds can be improved through distributed graphical processing unit (GPU) implementation; in fact, this has already been done for some applications (91). Additionally, the development of algorithms that provide a similar quality of results with fewer raw images or fewer computational steps or iterations will help speed processing and image reconstruction. An example is the use of the transport-of-intensity equation, as discussed in Section 2.6.
Finally, there is still a significant opportunity for the distribution of this technology in point-of-care or low-resource settings. Biomedical diagnostics in these areas is one of the primary motivations for the development of these compact, portable, and cost-effective lensless imaging and cytometry approaches. Although some field tests of lensless on-chip imaging have already been performed, much more work is required to test and evaluate these devices in the situations and settings that motivated them. Undoubtedly, this testing will reveal new areas for improvement that can be incorporated into later generations of lensless imaging, diagnostics, and cytometry devices.

**DISCLOSURE STATEMENT**
A.O. is the founder of a company, Holomic LLC, that commercializes computational imaging, sensing, and diagnostics technologies. A.O. and E.M. are coinventors of issued and pending patents on lensfree imaging and sensing techniques.

**LITERATURE CITED**

www.annualreviews.org • Lensless Imaging and Sensing

Changes may still occur before final publication online and in print.