Metasurface Holograms for Holographic Imaging

Weiwei Wan, Jie Gao,* and Xiaodong Yang*

As a revolutionary three-dimensional (3D) optical imaging technique, optical holography has attracted wide attention for its capability of recording both the amplitude and phase information of light scattered from objects. Holograms are designed to transform an incident wave into a desired arbitrary wavefront in the far field, which requires ultimate complex phase control in each hologram pixel. Conventional holograms shape the wavefront via the phase accumulation effect during the wave propagation through bulky optical elements, suffering issues of low-resolution imaging and high-order diffraction. Recently, metasurfaces, 2D metamaterials with ultrathin thickness, have emerged as an important platform to reproduce computer-generated holograms due to their advantages in manipulating light with well-controlled amplitude, phase, and polarization. In this article, the latest research progress in various types of metasurface holograms is reviewed from their design principles to versatile functional applications. At the end, more potential applications of metasurface holograms are discussed and some future research directions are also provided.

1. Introduction

Holography was originally invented by Dennis Gabor in 1948 for the purpose of improving electron microscope images.1 Due to the development of the laser in 1960s,2,3 holography found numerous optical applications for its powerful capability of arbitrary optical beam shaping and predesigned image recording and reconstruction. Holography is now recognized as a promising future technology for numerous applications in medicine, data storage, security, art, and entertainment. In contrast to traditional photography where only the intensity of light is stored but the information about the optical paths is lost, a hologram contains the complete information of the object beam so that real 3D image of an object can be rebuilt and the 3D image at different perspectives can be seen with the naked eye. Traditional holograms are created by recording the interference patterns between the light scattered by the object and the coherent reference beam.4,5 This technique produces an amplitude hologram, since the intensity of the interference pattern is directly recorded on the hologram. The conversion efficiency (power of the reconstructed beam versus that of the input beam) of amplitude holograms is very low, since much of the incident power is reflected or scattered. Fortunately, the interference intensity pattern can also be translated into phase variations to generate a phase hologram,6 which is usually preferred because it has a higher diffraction efficiency and can substantially increase the brightness of the reconstructed images. The phase profile in traditional phase-only hologram is controlled by etching different depths into a transparent substrate. Two-level binary phase holograms have been widely used because of their ease of fabrication, but they have a theoretical diffraction efficiency of only 40.5% and suffer the twin-image issue. Multilevel-phase holograms promise high diffraction efficiency and avoid the twin-image problem, but require expensive and complicated grayscale, variable-dose, or multistep lithography processes.7 Later on, computer-generated holography (CGH)8,9 was proposed to numerically calculate the phase and amplitude distributions at the hologram plane, which are then encoded into spatial light modulators (SLMs)10,11 or specific optical elements such as dielectric gratings12 and dielectric optical antennas.13 However, since conventional phase holograms rely on light propagation over distances much longer than the wavelength in order to accumulate enough phase variation for effective wavefront shaping, the optical elements for building the phase hologram will have large unitcell size and film thickness due to the limited refractive indices of natural materials. The smallest achievable hologram pixel size usually is as large as several micrometers (several times of the wavelength for visible light), which makes the hologram suffer the high-order diffraction, twin-image issue, and limited resolution.14 Therefore, compact holograms with subwavelength features in both the thickness and the pixel size are highly demanded, which will not only improve the quality of holographic imaging reconstruction but also reduce the hologram footprint for nanophotonic integration.

One significant breakthrough to bring new capabilities in confining light to subwavelength scale is plasmonics,15,16 where the coupling between electromagnetic waves and collective electron oscillations at metal–dielectric interfaces is studied. Due to the strong interaction with the incident light, plasmonic metamaterials provide more freedom in controlling light at visible and near-infrared frequencies due to their arbitrarily designed permittivity and permeability.17,18 However, the fabrication challenges at nanoscale and the absorption loss of bulk metamaterials at optical frequencies are usually
significant and inevitable, hindering their realistic applications. Recently, metasurfaces\textsuperscript{[20–28]} characterized as 2D metamaterials with ultrathin thickness, have been widely adopted for manipulating electromagnetic waves with reduced absorption loss and simplified fabrication process. Metasurfaces usually consist of flat optical element (meta-atom) arrays with spatially varying geometrical parameters and subwavelength separation, enabling us to tailor the spatial distributions of amplitude, phase, and polarization of light at subwavelength scale.\textsuperscript{[14,29–32]} Metasurfaces provide great flexibility in engineering the wavefront of light for versatile applications such as light bending,\textsuperscript{[20,21,33,34]} ultrathin flat lenses,\textsuperscript{[35–38]} optical vortex beam generation,\textsuperscript{[20,33,39–42]} and broadband quarter-wave or half-wave plates.\textsuperscript{[41,43–47]} The most complex and general wavefront shaping application is to create holographic images in the far-field. With the unique advantage of arbitrary wavefront control at subwavelength scale, the recently developed metasurface holograms enable the transformation of an incident wave into a desired wavefront profile in the far-field for reconstructing high-resolution holographic images without diffraction orders.\textsuperscript{[48–53]}

Here, we will review the recent research progress of metasurface holograms from the fundamental design principles to their versatile applications in holographic imaging with various functionalities. We will start by introducing the underlying mechanisms for the design of metasurface holograms, including two parts: First, the CGH algorithm used to generate the phase and amplitude map on the hologram plane; Second, several major types of meta-atoms used to control local phase and amplitude on metasurface holograms. Afterward, various types of high-performance metasurface holograms for achieving 3D imaging, high efficiency, full-color operation, and polarization selectivity will be discussed. At last, we will draw a brief conclusion and outlook for the metasurface hologram research.

2. CGH Algorithm

The production of metasurface holograms essentially involves two steps. The first step is to calculate the complex amplitude of the scattered wave from the object at the hologram plane by using CGH algorithm (which will be discussed in this section). The second step is to reproduce the calculated discrete phase and amplitude distributions by using the corresponding meta-atoms with various geometries as the pixels or building units for constructing the metasurface hologram (which will be discussed in Section 3).

CGH can produce wavefront profiles with any prescribed amplitude and phase distribution.\textsuperscript{[54]} It should be noted that, since metasurface holograms can directly record the phase distribution,\textsuperscript{[48,49,51]} it is not necessary to convert the phase information of an object into the variation of intensity by the interference between the scattered light from the object and a reference beam. Thus, one can directly calculate the light diffraction from the object on the hologram plane. The calculated 3D objects can be approximated as a collection of discrete point sources. The complex amplitude $U(x, y)$ on the hologram plane is calculated by superimposing the optical wavefront of all the point sources. The generated metasurface hologram can rebuild the object when it is illuminated by a plane wave without a reference beam. For simplicity, we assume that a virtual 2D object with its complex field of $U_0(x_0, y_0, z_0)$ is placed on the image plane $(x_0, y_0, z_0)$ with a distance $d = z_0 - z$ away from the hologram plane placed at $(x, y, z)$. Based on the Huygens–Fresnel principle, the diffraction patterns of each color component on the hologram plane can be described by
the Fresnel–Kirchhoff integral when the light incident angle is small[5,34,55]

\[
U(x,y,z) = \frac{1}{j\lambda} \int \int U_0(x_0, y_0, z_0) \exp\left(\frac{jkr}{r}\right) dx_0 dy_0
\]

(1)

where \( r = \sqrt{(x_0 - x)^2 + (y_0 - y)^2 + d^2} \) indicates the distance between two points on the object plane and the hologram plane. \( k = 2\pi/\lambda \) is the wavevector. If the reconstruction wavelength \( \lambda \) is typically many orders of magnitude smaller than the hologram dimension and the distance \( d \), we can apply the Fresnel approximation and get the Fresnel diffraction formula

\[
u(x,y,z) = \frac{\exp\left(\frac{jkd}{\lambda}\right)}{j\lambda} \exp\left[\frac{jk}{2d} (x^2 + y^2)\right] \int \int U_0(x_0, y_0, z_0) \exp\left\{-\frac{2\pi}{\lambda d} (x x_0 + y y_0)\right\} dx_0 dy_0
\]

\[
\times \frac{\exp\left(\frac{jkd}{\lambda}\right)}{j\lambda} \exp\left[\frac{jk}{2d} (x^2 + y^2)\right] 
\]

\[
\times F \left\{ U_0(x_0, y_0, z_0) \exp\left[\frac{jk}{2d} (x_0^2 + y_0^2)\right]\right\}
\]

(2)

where \( F \) represents the Fourier transform operation. \( U \) is the complex amplitude and can be written as \( U(x,y,z) = A(x,y,z) \exp(j\theta(x,y,z)) \). In order to apply the discrete Fourier transform of the complex amplitude, both the object plane and the hologram plane are usually divided into \( N \times N \) pixels. The resolution and the viewing angle of the reconstructed image can be controlled by designing the pixel size and the total dimension of the hologram, the working wavelength \( \lambda \), and the working distance \( d \). The amplitude and phase matrix can be calculated by using software such as Matlab. Some CGH example codes can be found in the books.[34,55]

If the phase-only hologram is needed, a random initial phase added to the point sources constituting the object can make the amplitude \( A(x,y,z) \) more uniform on the hologram plane. However, the reconstructed image by the above phase hologram is usually accompanied by speckle noise. The iterative Fourier transform algorithm (IFTA) can be used to overcome the speckle problem in holography. The IFTA was first proposed by Hirsch and co-workers.[8] Independently, Gerchberg and Saxton dealt with the phase-retrieval problem using a similar algorithm.[56] So the IFTA is also called the Gerchberg–Saxton algorithm, which is now widely used to generate the phase distribution on the hologram plane. We will not go to the details about the Gerchberg–Saxton algorithm which can be found in the references.[34,56,57]

3. Major Types of Meta-Atoms for the Control of Phase and Amplitude

The ideal metasurface holograms should be designed to simultaneously reproduce both the amplitude and phase information of the objects, which is calculated by the CGH algorithm. In the following, we introduce some representative types of meta-atoms which provide full control of the electromagnetic wave and can be used as the building blocks of metasurface holograms. The meta-atoms can be distinguished as three major different types based on the physical mechanisms to manipulate the phase variation under certain polarization states, including plasmonic dispersion for linearly polarized (LP) light, geometric effects with circularly polarized (CP) light, and Huygens’ principle with polarization insensitivity.

3.1. Meta-Atoms Based on Plasmonic Dispersion to Tune the Phase of LP Light

Due to the strong interaction between light and the localized surface plasmon resonances (LSPRs) in metallic antennas, both the phase and amplitude of the scattered wavefronts from the object can be tailored with a variety of parameters, such as material selection, antenna shape, dimension, and dielectric environment. LSPR is a dipole-like collective oscillation of electrons at the surface of metallic meta-atoms, and such collective electron motion can be described as a Lorentzian oscillator, being able to provide a phase tuning from 0 to \( \pi \) over the spectral width of the resonance.[58,59] For example, the phase response of a scattered (transmitted or reflected) beam at a fixed wavelength can be varied from 0 to \( \pi \) by changing the length of the nanorod antennas.[20,29,30] However, a complete phase control of the wavefront requires the entire 2\( \pi \) coverage, which can be realized by working with multiple resonances[20] or coupled resonances.[34]

3.1.1. Meta-Atoms Supporting Multiple Resonances

V-shaped nanoantennas, consisting of two arms of equal length joined together at a tunable angle, were first proposed and experimentally demonstrated to provide a full 2\( \pi \) phase coverage for the cross-polarized light (i.e., the output light is polarized in the direction perpendicular to the polarization of incident light) in mid-infrared wavelength range (Figure 1a).[20] Later on, the concept was extended to the near-infrared (NIR) range.[21] A V-shaped nanoantenna supports two plasmonic eigenmodes, symmetric and antisymmetric modes, which are excited by electric-field components along its diagonal axis and the orthogonal direction, respectively.[60] Under an arbitrary incident polarization, both the plasmonic eigenmodes can be excited but with substantially different amplitude and phase due to their distinctive resonance conditions. Thus, the double-resonance properties of V-shaped antennas provide a controllable phase delay and amplitude for the cross-polarized light by appropriately tuning the antenna geometrical parameters, such as the arm length, the angle between two arms, and the antenna orientation. The repeating unitcell highlighted in Figure 1a contains eight antennas which can provide equal scattering amplitudes and constant phase gradient of \( \pi/4 \) between the neighbors to cover the entire 2\( \pi \) phase. According to Babinet’s principle, the babinet-inverted V-shaped apertures have a similar optical response to their complementary structures of V-shaped antennas, but promise a higher signal-to-noise ratio by blocking the copolarized component for the transmitted light.[36,48] Nanoantennas with other kinds of geometries such
as C-shaped split-ring antennas can also be used to independently control the phase and amplitude of the cross-polarized light at the THz frequencies, determined by the opening angle and the orientation of antennas, respectively.\[61,62\] The antenna-based metasurfaces usually work over a wide wavelength range due to the low quality factor of antenna resonances.

3.1.2. Meta-Atoms Supporting Gap-Plasmon Modes

Beyond the antenna-based meta-atom designs, there have also been metasurfaces using reflect-arrays, which consist of metallic antennas separated from a metallic ground film with a thin dielectric spacer. Such metal–insulator–metal (MIM) nanostructures support magnetic resonances as gap modes inside the dielectric layer, relying on the strong coupling between the top antennas and their dipolar mirror images in the ground plane. Phase delays up to \(2\pi\) can be achieved and well controlled by changing the dimensions of the top antenna (Figure 1b).\[34,37,63,64\] By suppressing the transmission with the metallic ground plane, these three-layered reflect-arrays are able to provide extremely high reflection efficiency of 80% at the near-infrared frequencies.\[34\] Moreover, this approach has the advantage that the polarization state of the reflected light from the metasurface is the same as that of the incident light.

3.2. Meta-Atoms Based on Geometric Effects to Tune the Phase of CP Light

In Section 3.1, the phase tuning from the antenna is based on plasmonic resonances by controlling the antenna size or dimension. A completely different approach to manipulate the phase profile of circularly polarized light is to use the Pancharatnam–Berry (PB) phase (Figure 1c).\[65,66\] known also as geometric phase. PB phase originates from the geometric phase that accompanies polarization conversion rather than the optical path differences in conventional diffractive and refractive optical elements. As the pioneer in this research direction, Hasman and co-workers have first experimentally demonstrated the circularly polarized light conversion at a wavelength of 10.6 \(\mu\)m by using metallic or dielectric gratings.\[67,68\] By incorporating the PB phase control, they also used the dielectric gratings to achieve polarization-dependent focusing lens.\[69\]
Levy et al. have also used the rotated dielectric gratings to achieve CGHs at wavelengths of 1.55 and 10.6 µm. [12] Later on, the PB phase concept is extent to visible frequencies for various wavefront engineering applications. [33, 38, 49] The PB phase can be created by any anisotropic optical antennas. Assuming that the incident light is left-handed circular polarized (LCP) light with $E_i = E_{\text{LCP}} = \left[ \begin{array}{c} 1 \\ i \end{array} \right]$ (for right-handed circular polarized (RCP) light $E_i = E_{\text{RCP}} = \left[ \begin{array}{c} 1 \\ -i \end{array} \right]$), the scattered light from an anisotropic antenna with an orientation angle of $\alpha$ can be described as $|\alpha|^{74, 75}$

$$E_s = \left( \begin{array}{cc} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{array} \right) \left( \begin{array}{cc} t_e & 0 \\ 0 & t_o \end{array} \right) \left( \begin{array}{c} \cos \alpha \\ -\sin \alpha \end{array} \right)$$

$\times E_i = \frac{t_e}{2} E_{\text{LCP}} + \frac{t_o - t_e}{2} \exp(\pm i2\alpha) E_{\text{RCP}}$

where $t_e$ and $t_o$ are the light scattering coefficients for the two orthogonal linear polarization components of the incident light along the two axes of the scatter (antenna). Thus, the scattered light by a meta-atom antenna consists of two circular polarization states: one has the same handedness as the incident circularly polarized light (copolarization) without any phase delay (as shown by the first term of the above equation), and the other has the opposite handedness (cross-polarization) but with a phase delay of $\pm 2\alpha$ (as shown by the second term of the above equation). The sign “+” and “−” represent the phase delay for the incident LCP and RCP light, respectively (Figure 1c). This abrupt phase change occurring for circularly polarized light converted to its opposite helicity can be selected in experiments by using a quarter-wave plate and a linear polarizer. The key advantage of this approach lies in the sole and linear dependence of the phase delay $\Phi$ on the orientation angle $\alpha$ of the anisotropic antenna (i.e., $\Phi = \pm 2\alpha$). Thus, it can provide continuous phase control from 0 to $2\pi$ by simply rotating any kind of anisotropic antennas from 0 to $\pi$, such as nanorods, nanoslits, C-shaped, or U-shaped split-ring resonators. [39, 49, 51, 70] The PB phase generation is based on geometric effect and not subject to plasmonic dispersion, so that this approach provides ultrabroadband operation and it is robust against fabrication tolerances and variations of material properties. Meanwhile, the light scattering amplitude from the antenna resulted from the plasmonic resonances can be independently tuned by modifying the antenna size or dimension. By combining with the MIM reflect-array design, the PB-type phase control can achieve a reflection efficiency higher than 80%. [50]

3.3. Meta-Atoms Based on Huygens’ Principle

The polarization cross coupling efficiency of a single nonmagnetic plasmonic meta-atom (either antenna with LSPR or PB phase based anisotropic antenna) suffers from a theoretical limitation of 25%. [71] By introducing the magnetic response into the meta-atom, it is possible to achieve impedance matching between the metasurface and the free space to minimize the backward reflection and increase the coupling efficiency. Based on this idea, the Huygens’ metasurfaces [72, 73] are proposed to generate a unidirectional radiation field (with each metasurface pixel serving as a reflectionless secondary source) by controlling both the surface electric and magnetic polarizabilities $\epsilon_m$ and $\mu_m$ in order to match its impedance with the free space ($\sqrt{\epsilon_m/\mu_m} = n_m$). Certain equivalent electric and magnetic currents, which are proportional to $\epsilon_m$ and $\mu_m$, can be obtained by designing optical nanocircuit consisting of stacked multiple metallic and dielectric nanoblocks functionalized as inductors and capacitors, respectively (Figure 1d). [71, 73] The theoretical coupling efficiency of a Huygens’ metasurface designed by Alù and co-workers was predicted to be above 75% at the wavelength of 3 µm. [71] However, the measured efficiency of Huygens’ metasurfaces quickly drops below 20% at the NIR range ($\lambda = 1.2–2\mu$m). [73] due to the intrinsic metallic loss and weak magnetic response of the plasmonic meta-atoms. In addition to the metal–dielectric structures, all-dielectric meta-atoms can also be used to build the Huygens’ metasurfaces with low absorption loss (Figure 1e). It has been demonstrated that high refractive-index dielectric nanoparticles can simultaneously possess strong Mie-type electric and magnetic dipole responses [27, 29, 70]. The strong overlap and coupling between these resonances allows for full $2\pi$ control of phase of the incoming light. Silicon nanodisks were used by Decker et al. as the Huygens’ sources to constitute Huygens‘ metasurfaces with the transmission above 55% and a $2\pi$ phase coverage at the NIR frequencies. [76] Beyond the advantage of suppressed reflection, Huygens’ meta-atoms can be designed to be isotropic and hence insensitive to the light polarization state, [76, 77] providing the benefit to control of the phase of arbitrarily (linearly, circularly, or elliptically) polarized light.

It is worthy to mention that the meta-atoms suitable for building metasurface holograms are not limited to the above representative examples. More details about the different types of metasurfaces which can provide full control of the electromagnetic waves can be found in the review papers. [14, 29–32] Although the capability of fully producing the prescribed amplitude and phase distributions is desirable in designing metasurface holograms for yielding faithful optical images, other kinds of meta-atoms which can only provide the limited control of phase or amplitude can also be used as the building blocks of metasurface holograms by sacrificing the imaging quality. Once the amplitude and phase distributions are generated by CGH and the corresponding meta-atoms have been designed to reproduce the amplitude and phase profiles, the metasurface holograms can be fabricated by several commonly used surface-lithography techniques, such as photolithography, electron-beam lithography, focused ion beam milling, or nanoimprinting.

4. Various Types of Metasurface Holograms with Versatile Functionalities

Metasurface holograms can be divided into several categories depending on their phase/amplitude modulation, transmission/reflection output, and operating linear/circular polarization. Since human eyes are sensitive to different visible colors and metasurfaces can manipulate the polarization of light, here, we classify the metasurface holograms in terms of their responses of wavelength and polarization:
wavelength-insensitive (single-color) metasurface holograms, wavelength-multiplexed (multicolor or full-color) metasurface holograms, polarization-insensitive metasurface holograms, and polarization-multiplexed metasurface holograms (with LP and CP light).

4.1. Wavelength Insensitive (Single-Color) Metasurface Holograms

Due to the broad bandwidth of the plasmonic resonances and the PB phase generation, metasurface holograms without wavelength-multiplexing can usually reconstruct one image with the color determined by the wavelength of the laser beam. This type of metasurface holograms cannot distinguish different colors in one image, thus they are insensitive to the wavelength and can be called as single-color metasurface holograms. We will introduce some typical single-color metasurface holograms, which are composed of the meta-atoms mentioned in Section 3, exhibiting good performance such as high resolution, 3D imaging, or high efficiency.

Complementary V-shaped nanoantennas perforated in a 30 nm gold film on a glass substrate have been used to reconstruct a holographic image at a visible wavelength of 676 nm (Figure 2a). The babinet-inverted V-shaped aperture is advantageous in the flexibility of phase/amplitude control by changing their arm lengths and splitting angles, with which the eight-level phase modulation from 0 to 2π and the two-level amplitude modulation as 0 and 1 have been experimentally demonstrated. High imaging resolution is achieved because of its smallest pixel size of 150 × 150 nm² in the existing metasurface holograms. This transmission-type metasurface hologram based on plasmonic resonances has an overall efficiency of 10% (the fraction of the transmitted energy of the cross-polarized light that contributes into the holographic image). Different from metasurface holograms under linear polarization, PB phase based metasurface holograms work under circular polarizations and encode the desired phase profiles into orientation-rotated anisotropic nanoantenna array. 3D holographic images with a wide field of view estimated as 40° have been experimentally demonstrated by designing the PB phase profile with the array of gold nanorods with various orientation angles on a glass substrate (Figure 2b). Single-layered plasmonic metasurfaces usually suffer low efficiency, with most of the energy lost in the reflection and cross-polarization conversion.

One approach to address the low transmission efficiency of the single-layered plasmonic metasurface hologram is to use MIM reflection configuration. Zhang and co-workers combined the PB phase control of plasmonic metasurfaces with MIM reflect-arrays to yield a high-efficiency reflective metasurface hologram with 16-level-phase modulation (Figure 2c). Similar to the half-wave plate, the reflection along the long axis and the short axis of the nanorod antenna is designed to have a phase difference of π for fully converting a circularly polarized beam into the oppositely polarized one. By blocking the transmission with a ground metallic film and increasing the polarization conversion efficiency, such reflection-type phase-only metasurface hologram can achieve diffraction efficiencies of 80% at 825 nm and over 50% in a broad bandwidth of 630–1050 nm, making the metasurface holograms more practical for applications. Another approach to increase the efficiency of the transmission-type metasurface hologram is to eliminate the reflection by using the Huygens’ principle. The dielectric Huygens’ metasurface holograms consisting of silicon nanopillars with 36 different radii have been experimentally demonstrated to transmit over 90% of light with a diffraction efficiency over 99% at the wavelength of 1600 nm (Figure 2d).

4.2. Multicolor Metasurface Holograms

Different from single-color metasurface holograms, multicolor metasurface holograms are able to reconstruct multiple images with different colors at the same time. However, the amplitude and phase response of plasmonic resonator is usually dispersive and broadband, which make the design of multicolor metasurface hologram complicated. One effective approach is to design the plasmonic resonances with narrow bandwidths, which allows multiple meta-atoms to resonate at separated wavelengths. Thus, a super unitcell consisting of multiple meta-atoms with different resonant wavelengths can act as one pixel with the capability of controlling multiple colors independently. Tsai and co-workers used the super unitcell concept to design binary-phase holograms working for three primary colors (red, green, and blue) (Figure 3a). They designed three pairs of meta-atoms based on the MIM configuration consisting of a top layer of aluminum nanorods, a dielectric spacer layer, and a bottom aluminum ground film, where the phase modulation is realized by varying the nanorod length. These three pairs of meta-atoms can independently provide a binary-phase (0 and π) control for three primary colors, respectively. In order to avoid the reflection amplitude variations arising from the different size of the nanorods, each super unitcell is designed to contain four meta-atoms (two for red light, one for green light, and one for blue light) to compensate the low reflection amplitude of the red light which are almost the half of those of the green and blue light. By combining narrow-bandwidth plasmonic resonance as the wavelength filter and the PB phase as the phase modulation, Choudhury et al. used three silver nanoslits with different sizes and orientations to independently control the transmitted amplitude and phase of three primary colors, respectively, for rebuilding three-color holographic images (Figure 3b). Besides the three-color phase-modulation metasurface holograms, the two-color amplitude-modulation metasurface holograms have also been experimentally demonstrated by using the scattering of metallic nanoparticles and the transmission of the fishnet metamaterials. Montelongo et al. designed four basic pixels consisting of silver nanoparticles and the transmission of the fishnet metamaterials. Montelongo et al. designed four basic pixels consisting of silver nanoparticles and the transmission of the fishnet metamaterials. Montelongo et al. designed four basic pixels consisting of silver nanoparticles and the transmission of the fishnet metamaterials. Montelongo et al. designed four basic pixels consisting of silver nanoparticles and the transmission of the fishnet metamaterials.
4.3. Full-Color Metasurface Holograms

Above three-color metasurface holograms certainly have the potential for full-color holographic applications by using the color mixing. However, the super unitcell design containing nanoantennas with different sizes inevitably increases the pixel size. On the other hand, in most of the recent demonstrations for full-color holograms, only the phase information has been engineered while the amplitude information is lost, resulting in an inevitable limitation in obtaining high-resolution images and further color mixing. Wan et al. have demonstrated full-color ultrathin plasmonic metasurface holograms with both the amplitude and phase modulations for reconstructing 2D and 3D holographic images with high resolution (Figure 4a).[51]
The full-color hologram is generated by superimposing three CGH-holograms which are calculated separately for the red, green, and blue components and encoded with an additional phase shift for each color. In the reconstruction process, the illumination of three lasers with different tilted incident angles is used to decode three color components to generate the desired full-color images. The final full-color holograms have only a single metallic nanoslit in one pixel to independently control the transmitted amplitude and phase of three primary colors by tuning the nanoslit length and orientation (by using the PB phase), respectively, where the two-level amplitude and eight-level phase modulations have been demonstrated. This full-color plasmonic metasurface hologram with subwavelength pixel size can produce not only three primary colors but also their secondary colors (cyan, magenta, yellow, and white), which is promising for generating high-resolution full-color holographic images without other diffraction orders.

The full-color metasurface holograms have also been demonstrated by other approaches. Ozaki et al. reconstructed 3D color holographic images based on surface plasmons (Figure 4b), where each color is reconstructed by satisfying the resonance condition of surface plasmons at the individual wavelength. Such surface plasmon hologram is first recorded on a photore sist by the interference patterns between the scattered light field from the object and the reference laser beams of three colors, and then a thin silver film is coated on the photoresist hologram to generate the plasmonic gratings, which can convert the surface plasmons at different wavelengths into the radiative light field during the reconstruction process. This surface plasmon hologram has the advantage of rebuilding color images with white light illumination, because the surface plasmons at different wavelengths can be selectively excited by the white light with the corresponding incident angles. The full-color metasurface holograms have also been demonstrated by using dielectric metasurfaces. By combining the supercell design with the PB phase, Wang et al. used three types of silicon nanoblocks with various dimensions and orientation angles as three independent channels in one super unitcell to manipulate the phase profiles for the red, green, and blue light (Figure 4c). Li et al. demonstrated that graphene oxides can also be used to record the phase profiles of full-color objects through an athermal photoreduction process induced by a single femtosecond pulse (Figure 4d).

4.4. Polarization-Insensitive Metasurface Holograms

Since meta-atoms are usually anisotropic, most of the metasurface holograms are polarization sensitive so that the phase control of linearly or circularly polarized light is typically...
realized. The Huygens’ meta-atoms rely on the overlap of the electric and magnetic resonances in the high-index dielectric nanoparticles or the stacked plasmonic structures to provide the complete phase control, and they can be designed to be isotropic for exhibiting polarization-insensitive properties. Kivshar and co-workers employed isotropic silicon nanodisks with different lattice periodicities to achieve a four-level phase-modulation metasurface hologram with 82% transmittance efficiency and 40% imaging efficiency at the NIR frequency (Figure 5a).[77] By introducing more phase levels provided by 36 different silicon nanodisks with various radii, the diffraction efficiency can be further improved over 99%. [78] Zhao et al. also take advantage of the high efficiency of Huygens’ surface to realize the polarization-independent metasurface hologram at the wavelength range of 670–900 nm with transmission efficiency up to 86% and optical efficiency of 23.6%. [85] Besides the isotropic dielectric nanodisks, Huang et al. used a photon sieve with a large number of isotropic nanoholes randomly distributed in a chromium film to realize a polarization-independence hologram with diffraction efficiency of 47% (Figure 5b). [86] These photon sieves control the wavefront of light based on diffraction and interference, which can be produced using techniques similar to those for making the zone plates. [87]

4.5. Linear or Circular Polarization-Multiplexed Metasurface Holograms

The polarization-multiplexed metasurface holograms are sensitive to the incident polarization state and can reconstruct different holographic images by switch the incident polarization. Tsai and co-workers used the MIM reflection-type configuration (cross-shaped gold nanorods on top of a dielectric layer on a gold ground mirror) to achieve a four-level phase-modulation metasurface hologram for reconstructing switchable dual images controlled by the incident linear polarizations (Figure 5c). [88] By introducing the gap-plasmon mode, the phase of the reflected wave can be tuned independently by varying the nanorod length. The electric dipole resonance of a single nanorod is selectively excited by the incident light with the linear polarization along the nanorod direction. Therefore, two holograms consisting of nanorod arrays are turned 90° to each other to form one combined hologram with the cross-shaped nanorod arrays. Two distinct images can be reconstructed under the orthogonal linearly polarized light, with an experimental efficiency of 18% for the 780 nm illumination. Montelongo et al. used the same methodology of selectively exciting the vertically and horizontally orientated silver nanorods with identical sizes in a “L” shape to achieve a binary-ambiguity metasurface
hologram for generating linear polarization-switchable images (Figure 5d).

In addition, the polarization multiplexed metasurface holograms can also be designed to work under circularly polarized light, which is also called as helicity-multiplexed metasurface holograms. Zhang and co-workers used the three-layer MIM design combining with the PB phase control of the silver nanorods atop to realize ultrabroad-bandwidth and high-efficiency helicity-multiplexed metasurface holograms with efficiencies of 59.2% at 860 nm and over 40% in the spectral range of 475–1100 nm (Figure 5e). The identical silver nanorods with various orientation angles can provide not only almost continuous phase profile with 16 levels in the entire 2π range, but also uniform reflection amplitude by eliminating the unintended amplitude variations arisen from the different size of the top nanorods. Two sets of hologram patterns operating with the opposite incident helicities were combined together with interlaced displacement in the final metasurface hologram. Two symmetrically distributed off-axis images can be interchangeable by controlling the helicity of the input light. Capasso and co-workers also used the dielectric metasurfaces consisting of two sets of rotated silicon nanoblocks to project different images depending on the handedness of the incident beam (Figure 5f).

4.6. Wavelength- and Polarization-Multiplexed Nonlinear Metasurface Holograms

It is worthy to note that Zhang and co-workers have experimentally demonstrated that both the wavelength and circular polarization can be simultaneously multiplexed into one nonlinear metasurface hologram (Figure 6a). The nonlinear metasurface hologram consists of gold split ring resonators as meta-atoms with strong polarization properties in the linear regime as well as high second-harmonic generation efficiency in the nonlinear regime. The metasurface hologram is realized by utilizing the PB phase change operating in both the
linear and nonlinear regimes at the same time. Under the excitation of a CP incident beam with spin state $\sigma$ ($\sigma = \pm 1$ for LCP and RCP), the PB phase for the transmitted fundamental beam (around 1120 nm) with the opposite spin $-\sigma$ is $2\sigma\alpha$, while the nonlinear PB phase for the transmitted second-harmonic generation beam (around 560 nm) with the same spin $\sigma$ and the opposite spin $-\sigma$ is $\sigma\alpha$ and $3\sigma\alpha$, respectively. As a result, three circular polarization-encoded linear and second-harmonic generation channels are utilized to generate three independent holographic images of the letters “X,” “R,” and “L” simultaneously (Figure 6a), by selecting the wavelength and spin of the transmitted light without changing the incident CP light and the metasurface. Thus, the nonlinear metasurface holograms allow the reconstruction of multiple target holographic images carried independently by crosstalk-free postselective channels of the fundamental and harmonic generation waves with different spins.\cite{91} In addition, Almeida et al. have demonstrated polarization-sensitive multilayered nonlinear metasurface holograms with the third-harmonic generation, by which a NIR incoming beam (at 1266 nm) is converted into the switchable blue-color (at 422 nm) holographic images selected by the incident linear polarization (Figure 6b).\cite{92} Linearly polarizable V-shaped gold antennas with various arm lengths and angles are used as the meta-atoms with tunable plasmonic resonances for providing variable phase shifts $\varphi$ on the incoming fundamental beam. Then the nonlinear metasurface hologram is designed with the nonlinear phase shift $3\varphi$ of the third-harmonic generation signal. Two-layer nonlinear metasurface hologram was made with a separate phase hologram embedded in each metasurface layer to yield the desired far-field image for the incident vertical or horizontal linear polarization. Figure 6b shows the generated holographic images of happy and sad smiley faces with each image recreated only by the certain linearly polarized input beam. Compared to the linear metasurface holograms, the nonlinear metasurface holograms working at the harmonic frequencies remove the image background noise at the fundamental wavelength. In general, the nonlinear optical processes (second- or third-harmonic generation) provide more multiplexing channels in building metasurface holograms for reconstructing multiple complex images, showing great potential in future applications.

5. Conclusions and Outlook

Metasurfaces provide a good platform for realizing the holograms with high performance due to the capability of fully manipulating electromagnetic wave in the amplitude, phase, and polarization. We have reviewed a few representative types of meta-atoms suitable as the building blocks of the metasurface holograms to control local phase and amplitude. Various types of metasurface holograms with high performance, such as 3D imaging, high resolution, high efficiency, polarization selectivity, and full-color operation, have been summarized from the perspective of achieved functionalities. Besides the holographic imaging applications, the holography principle has also been used as a powerful tool to design either plasmonic interfaces to excite the desirable surface plasmon waves or metasurfaces for free-space beam wavefront shaping with any desirable field distributions, such as laser emission collimation,\cite{93,94} light orbital angular momentum detection,\cite{95,96} optical vortex and vector beam generation,\cite{24,10,97,98} nondiffraction plasmonic

Figure 6. Nonlinear metasurface holograms. a) Spin- and wavelength-multiplexed nonlinear metasurface holograms. Linear hologram for reconstruction of the letter “X” at NIR wavelengths around 1120 nm and nonlinear holograms for generating second-harmonic generation holographic images of the letters “R” and “L” encoded in opposite spins around 560 nm. Reproduced with permission.\cite{91} Copyright 2016, NPG. b) Two-layer polarization-switchable nonlinear metasurface holograms based on linearly polarizable V-shaped gold antennas. Two third-harmonic generation holographic images of happy and sad smiley faces at 422 nm with each image generated by one hologram layer under vertically or horizontally polarized incident fundamental beam at 1266 nm, respectively. Reproduced with permission.\cite{92} Copyright 2016, NPG.
beam launching\textsuperscript{[99,100]} and polarization-sensitive directional couplers and beam splitters\textsuperscript{[23,101]} More detailed applications of holographic optical metasurfaces can be found in the review by Genevet and Capasso.\textsuperscript{[102]}

Although metasurfaces have vastly extended the application of holograms, the efficiency of the single-layer transmission-type plasmonic metasurface hologram is usually very low (<10\%). High diffraction efficiencies (up to 80\%) of plasmonic metasurface holograms have been limited to the reflection-type operation and working under circularly polarized incident light. The absorption loss in plasmonic metasurfaces as well as the reflection issue and low polarization-conversion efficiency pose a fundamental obstacle for achieving high-transmission metasurface holograms. It seems that the above limitations of plasmonic metasurfaces can be partially overcome by using all-dielectric Huygens’ metasurfaces made of silicon nanodisks which have smaller losses and less reflection at the NIR frequencies.\textsuperscript{[78]} The PB phase based metasurfaces consisting of silicon nanoposts have been used for making holograms for visible spectrum, but still with certain optical absorption losses.\textsuperscript{[53,103]}

As one promising solution, Capasso and co-workers have recently developed all-dielectric PB phase based metasphere holograms made of titanium oxide nanofins to provide significantly increased conversion efficiencies for generating holographic images with high transmission and broad bandwidth at the visible frequencies.\textsuperscript{[90,104,105]} Another recent interesting method with the combination of PB phase and propagation (or dynamical) phase in tandem in the meta-atoms allows for the imposition of arbitrary phase profiles on any two orthogonal polarization states (linear, circular, or elliptical),\textsuperscript{[106]} resulting in polarization-switchable chiral holograms where a single metasurface encodes two independent hologram phase profiles for each circular polarization.

Besides the conversion efficiency issue, the broadband operation is another key factor that determines the performance of metasurface holograms. One can suppress the dispersion by designing antenna resonances with low quality factors or even completely avoid the dispersion by using the dispersion-free PB phase response that relies only on the geometric effect. On the other hand, the dispersion of each unitcell can also be used to generate a sharp resonance as the wavelength filter for wavelength-multiplexing configuration. Thus, one can eliminate or utilize the dispersion of metasurface holograms for satisfying different bandwidth requirements.

The resolution of a metasurface hologram is determined by the pixel size of each meta-atom unitcell. The small separation between neighboring meta-atoms will introduce not only high resolution for the hologram but also the near-field interaction between meta-atoms. Such cross-talk will result in phase and amplitude distortion and hence degrade the quality of imaging reconstruction. Controlling the meta-atoms sufficiently well separated is an efficient approach to suppress the cross-talk. For the polarization- and wavelength-multiplexing holograms, the cross-talk between different polarization or wavelength channels can also be minimized by placing meta-atoms off-axis.\textsuperscript{[81,89]} The nonlinear holograms provide different wavelength channels, which can completely eliminate the cross-talk issue for multiplexing and multidimensional holographic applications.

Metasurface holograms consisting of 2D arrays of plasmonic or dielectric nanostructures in a thin film can be fabricated by commonly used surface-lithography techniques, such as photolithography, electron-beam lithography, focused ion beam milling, or nanoimprinting. Photolithography is widely used in making semiconductor integrated circuits because of its high throughput at the micrometer scale. However, the sub 100 nm feature requires advantageous and expensive deep-ultraviolet photolithography which is typically available for commercial products. Compared with photolithography, electrobeam lithography and focused-ion-beam lithography have high cost and low throughput, but they are straightforward and precise nanofabrication tools to make metasurface hologram prototypes in small areas, which are excellent for research purpose. Nanoimprint lithography is a low-cost nanopatterning technology for the mass production of metasurface holograms with nanoscale feature sizes once the imprinting mold is created. The imprinting mold can be used for many times to replicate the metasurface holograms into imprinting resist.

Usually, the optical properties of the existing metasurface holograms are fixed after the device fabrication and lack of tunability. As their counterparts, SLMs with bulky sizes such as liquid crystal displays and digital micromirror devices have the flexibilities to dynamically modulate the phase and amplitude of light, leading to the real-time holographic display applications. By incorporating different approaches that can tune the phase responses of metasurface pixels in real time such as thermo-optic effect,\textsuperscript{[107,108]} phase-change materials,\textsuperscript{[107,108]} electronic control,\textsuperscript{[109,110]} mechanical stain,\textsuperscript{[111–113]} and free-carrier injection,\textsuperscript{[114,115]} dynamically tunable metasurface holograms with varying functionalities on demand can be demonstrated in the near future. So far, the wavelength- and polarization-multiplexed metasurface holograms have already been demonstrated. More channels to store the optical information can be further added into the metasurface holograms by utilizing new methods such as multiple nonlinear optical effects, orbital angular momentum multiplexing,\textsuperscript{[116]} and quantum phenomena of single photon,\textsuperscript{[117]} for providing high-capacity imaging and information processing.

Metasurface holograms have been developed quickly in the past few years with versatile demonstrated functionalities. There is still big room for the advancement of metasurface holograms with high performance, dynamical tunability, multiple-channel encoding, and on-chip or fiber integration, which will open more expected opportunities in the future applications of high-capacity optical imaging and information processing, complex beam conversion, data storage and encryption, anti-counterfeiting, optical trapping, and integrated photonics.

Acknowledgements

The authors acknowledge support from the National Science Foundation under Grant No. ECCS-1653032 and DMR-1552871, and the Office of Naval Research under Grant No. N00014-16-1-2408.

Conflict of Interest

The authors declare no conflict of interest.
Keywords
holo&copy;grphy, imaging, metasurfaces


[78] L. Wang, S. Kruk, H. Tang, T. Li, I. Kravchenko, D. N. Neshev, Y. S. Kivshar, Optica 2013, 3, 1504.


