From Metasurfaces to Compact Optical Metasystems

Thesis by
MohammadSadegh Faraji-Dana

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy

Caltech

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2020
Defended June 5, 2020
To my love, family and all who have taught me
Almost four years ago (in September 2016) I joined Caltech and my Ph.D. academic journey started. It is quite unbelievable to me now how my perspective to life, science, society and many others have been greatly shaped throughout these years. In fact, the unique experiences I gained at Caltech played a significant role in developing my personal skills, attitudes as well as providing me with potential career opportunities. I have been blessed to have the presence of amazing people around me, especially my friends, mentors, colleagues, and on top of the list my Ph.D. advisor. While, it is not possible to name all of them here, I would like to take this opportunity to express my gratitude to all of those who helped and supported me throughout these years of my graduate experience.

I have been fortunate to have Professor Andrei Faraon as my Ph.D. supervisor. He has been not only an always supportive advisor but also a true role model of a great scientist, who has been providing his students the freedom and insight to explore and investigate new ideas in their research. He was truly inspiring to me both personally and also scientifically and I am greatly thankful of him for his passionate and supportive personality as well as his patience and positive attitude at the times of difficulties and failures. I am not able to imagine a better advisor than him for my Ph.D. studies.

Many of the faculty members inside and outside Caltech have helped me by promoting the sense of curiosity and questioning in these years; in specific I appreciate my candidacy and thesis defense committee members, Professors Alireza Marandi, Yu-Chong Tai, Lihong Wang, and Amnon Yariv for their valuable time and thoughtful pieces of advice. I also take this opportunity to thank our Caltech collaborators, Professors Dimitry Mawet, and Changhuei Yang for the insightful discussions and collaborations on various ideas.

I am grateful to all Caltech staffs for their kind assistance and for providing a friendly and supportive environment for study and research, especially Cecilia Gamboa, and Tanya Owen for their tireless academic support. I would like to also thank Christine Jenstad, Carol Sosnowski, and Jennifer Blankenship for organizing events, seminars and talks. In addition, I would like to thank Dr. Kate McAnulty, Natalie Gilmore, and Angelica Medina-Cuevas from the Graduate studies office for their endless support. Furthermore, I would like to thank Tess Legaspi from the Registrar’s Office, Laura
Flower Kim, Daniel Yoder, and Ilana Smith from the International Student Office for being always available and kindly assisting me in any aspect they could.

It should be mentioned that almost none of the work in this thesis would have been possible without the great support received from the Kavil Nanoscience Institute (KNI) at Caltech. I am extremely grateful to all KNI staffs especially Dr. Guy DeRose, Dr. Matthew Sullivan Hunt, Bert Mendoza, Alex Wertheim, Nathan Lee and Tiffany Kimoto. I would like to also thank our collaborators from Samsung Electronics which supported this work, especially Dr. Seunghoon Han and Dr. Duhyun Lee for their fruitful ideas and discussions.

I owe a debt of gratitude to all past and present Faraon group members, who have been always kind and helpful, and contributed in creating a friendly, supportive and collaborative ambient in the group. In specific, I would like to thank Dr. Amir Arbabi, currently an assistant professor at UMass Amherst, for his insightful ideas and persistent guidance and support. I am extremely grateful to Dr. Ehsan Arababi and Dr. Mahsa Kamali for mentoring me in the lab and the clean-room and above all of that for being my true friends in these years. I am also greatly thankful to Hyounghan Kwon for helping me a lot with my research and for all the great discussions and collaborations we had. I like to express my gratitude to Conner Ballew, Tianzhe Zheng, Greg Roberts and Philip Camayd-Munoz for what I have learned from them in our discussions. I would also like to thank Ioana Craiciu, Evan Miyazono, John Bartholomew, Tian Zhong, Jonathan Kindem, Chuting Wang, Jake Rochman, Andrei Ruskuc, Daniel Riedel, Ding Zhong, Mi Lei, and Tian Xie for always being there to help and whom I was able to count on whenever needed.

I owe a tremendous gratitude to all people who played a significant role in my scientific upbringing and fostering my passion for science and engineering in these years. I would like to specifically thank my master’s degree advisor Professor Safieddin Safavi-Naeini to whom I owe my knowledge of electromagnetics and waves. In addition, I would like to thank professors Hamed Majedi, Nasser Masoumi, Karim Mohammadpour-Aghdam, and Jalil Rashed-Mohassel, who were my sources of knowledge and inspiration in many different aspects.

I have been very fortunate to be raised in an academic family were engineering and science were always highly valued. I have been really blessed for having the unconditional support of my family throughout my graduate studies years. I like to take this opportunity to thank my parents, and also my sister for being extremely supportive of me in any way they could in spite of all difficulties caused by being
thousands of miles away from each other. I truly appreciate their help, support and mainly their patience throughout these years and I hope sometime in the near future my family could be altogether. That being said I was blessed to have several good friends in the Los Angeles and North America who have been always the source of joy and happiness in my life. While it is impossible to name them all here, I would like to specifically thank Mohsen, Fatemeh, Ebrahim, Mahdi, Bahareh, Hadi, Alireza, Ali, Mohammad, Mahmood, Fariborz, Ehsan, Mahsa, Reza, Parham, Hassan, Massoud, Amir, Ghazaleh, Hamidreza, Kousha, Benyamin, and Hossein.

My PhD studies at Caltech coincided with meeting my companion in life, Azin. As we were separated from each other by thousands of miles, my long-lasting academic journey continued without having her on my side. Today, I am very happy as I was able to follow both my scientific dream and also having her as my life companion. This would have been absolutely impossible without her exceptional patience, loyalty and devotion. I am afraid I will never be able to properly thank for her kindness, understanding and support.

And finally, my journey of life has been always full of unprecedented experiences, unparalleled opportunities, new relationships and people which I have nearly had minimal control on many of them. I am deeply grateful to all who made these happen to me, and above all of those, I am thankful to God for his everlasting blessings.

MohammadSadegh Faraji-Dana
Pasadena, California
June 11, 2020
ABSTRACT

Optical metasurfaces are a class of ultra-thin diffractive optical elements, which can control different properties of light such as amplitude, phase, polarization and direction at various wavelengths. The compatibility of optical metasurfaces with standard micro- and nano-fabrication processes makes them highly-suited for realization of compact and planar form optical devices and systems. In addition, optical metasurfaces have achieved unique and unprecedented functionalities not possible by conventional diffractive or refractive optical elements. In this thesis, after a short review on the history and state of the art optical metasurfaces, I will discuss the systems consisting of optical metasurfaces, called optical meta-systems, which allow for implementations of complicated optical functions, such as wide field of view imaging and projection, tunable cameras, retro-reflection, phase-imaging, multi-color imaging, etc. Thereafter, the concept of folded metasurface optics is introduced and a compact folded metasurface spectrometer is showcased to demonstrate how the folded meta-systems can be designed, fabricated and practically utilized for real-life applications. Furthermore, different approaches for implementation of miniaturized hyperspectral imagers are investigated, among which the folded metasurface optics and a computational scheme using a random metasurface mask will be highlighted. Other potentials of optical metasurfaces achieved by the employment of optimization techniques to improve their multi-functional performances, as well as example applications in realizing optical vortex cornographs are studied. Finally, I will conclude the dissertation with an outlook on further applications of optical metasurfaces, where they can surpass the performance of current optical devices and systems and what limitations are still to be overcome before we can expect their wide-spread applications in our daily life.
Thesis supervisor:

- Andrei Faraon
  Professor of Applied Physics and Electrical Engineering

Thesis committee:

- Andrei Faraon
  Professor of Applied Physics and Electrical Engineering

- Alireza Marandi (chair)
  Assistant Professor of Electrical Engineering and Applied Physics

- Yu-Chong Tai
  Anna L. Rosen Professor of Electrical Engineering & Medical Engineering

- Lihong Wang
  Bren Professor of Medical Engineering and Electrical Engineering

- Amnon Yariv
  Martin and Eileen Summerfield Professor of Applied Physics and Electrical Engineering
PUBLISHED CONTENT AND CONTRIBUTIONS


# TABLE OF CONTENTS

Acknowledgements ................................................. iv
Abstract ............................................................. vii
Published Content and Contributions ........................... ix
Table of Contents .................................................. ix
List of Illustrations ............................................... xi
Abbreviations ....................................................... xiii

Chapter I: Introduction ............................................ 1
  1.1 A history of optical metasurfaces ........................... 1
  1.2 State of the art optical metasurface devices ............... 4
  1.3 Optical meta-systems ....................................... 7
  1.4 Thesis outline ............................................. 10

Chapter II: Folded optical meta-systems ......................... 11
  2.1 Compact Folded metasurface spectrometer ................. 11
  2.2 Appendix: Additional information and discussion for the metasurface spectrometer ........................................ 23
  2.3 Appendix: Additional figures and tables for the metasurface spectrometer ................................. 28

Chapter III: Miniaturized hyperspectral imagers ............... 37
  3.1 Hyperspectral imager with folded metasurface optics ..... 37
  3.2 Random phase mask metasurface hyperspectral imager .... 47
  3.3 Appendix: Additional information and discussion for folded metasurface hyperspectral imager ....................... 54
  3.4 Appendix: Additional information and discussion for random phase mask metasurface hyperspectral imager ............ 64

Chapter IV: Optical metasurfaces for coronographs ............ 67
  4.1 Inverse design of high efficiency optical metasurfaces .. 67
  4.2 Optical vortex coronographs ............................... 70
  4.3 Appendix: Additional information and discussion for inverse design of high efficiency optical metasurfaces ............. 73

Chapter V: Conclusion and outlook ................................ 75

Bibliography ....................................................... 77
# LIST OF ILLUSTRATIONS

<table>
<thead>
<tr>
<th>Number</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 A short history of optical metasurfaces</td>
<td>3</td>
</tr>
<tr>
<td>1.2 State of the art optical metasurfaces</td>
<td>5</td>
</tr>
<tr>
<td>1.3 Optical meta-systems</td>
<td>8</td>
</tr>
<tr>
<td>2.1 Schematics of a conventional and a folded metasurface spectrometer</td>
<td>13</td>
</tr>
<tr>
<td>2.2 Ray-optics design of the spectrometer</td>
<td>15</td>
</tr>
<tr>
<td>2.3 Metasurface structure and design graphs</td>
<td>16</td>
</tr>
<tr>
<td>2.4 Spectrometer characterization results</td>
<td>18</td>
</tr>
<tr>
<td>2.5 Sample spectrum measurement results</td>
<td>20</td>
</tr>
<tr>
<td>2.A1 Nano-post simulation for folded spectrometer</td>
<td>28</td>
</tr>
<tr>
<td>2.A2 Reflection phase variation versus wavelength</td>
<td>29</td>
</tr>
<tr>
<td>2.A3 Grating design curves and deflection efficiencies</td>
<td>29</td>
</tr>
<tr>
<td>2.A4 Fabrication process for spectrometer</td>
<td>30</td>
</tr>
<tr>
<td>2.A5 Spectrometer measurement setups</td>
<td>30</td>
</tr>
<tr>
<td>2.A6 Focal plane intensity profiles of the spectrometer</td>
<td>31</td>
</tr>
<tr>
<td>2.A7 Intensity distribution profiles on logarithmic scales</td>
<td>32</td>
</tr>
<tr>
<td>2.A8 Measured spectral resolution versus wavelength</td>
<td>32</td>
</tr>
<tr>
<td>2.A9 Angular response of the spectrometer</td>
<td>33</td>
</tr>
<tr>
<td>2.A10 High-throughput spectrometer design</td>
<td>34</td>
</tr>
<tr>
<td>3.1 Concept of a push-broom folded metasurface HSI</td>
<td>40</td>
</tr>
<tr>
<td>3.2 Ray-optics design and images of the fabricated devices</td>
<td>41</td>
</tr>
<tr>
<td>3.3 Characterization results of the folded metasurface HSI</td>
<td>44</td>
</tr>
<tr>
<td>3.4 Hyperspectral imaging results using the folded metasurface HSI</td>
<td>46</td>
</tr>
<tr>
<td>3.5 Hyperspectral imaging using a random metasurface phase mask</td>
<td>49</td>
</tr>
<tr>
<td>3.6 Optical properties of a uniform distribution phase mask</td>
<td>50</td>
</tr>
<tr>
<td>3.7 Optical properties of a phase mask designed using the GS algorithm</td>
<td>51</td>
</tr>
<tr>
<td>3.8 Spatial and spectral correlation of designed phase mask</td>
<td>52</td>
</tr>
<tr>
<td>3.9 Spectral content of the object used for hyperspectral imaging</td>
<td>52</td>
</tr>
<tr>
<td>3.10 Object reconstruction at different wavelengths</td>
<td>53</td>
</tr>
<tr>
<td>3.A1 Optimized phase profiles and simulated PSFs for the HSI</td>
<td>55</td>
</tr>
<tr>
<td>3.A2 Unit cells and curves showing designed nano-posts dimensions</td>
<td>57</td>
</tr>
<tr>
<td>3.A3 Reflection phase variation versus incidence angle</td>
<td>58</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------------------------------------------------------</td>
</tr>
<tr>
<td>3.A4</td>
<td>HSI measurement and characterization setups</td>
</tr>
<tr>
<td>3.A5</td>
<td>Characterization results of the metasurface HSI</td>
</tr>
<tr>
<td>3.A6</td>
<td>Measured efficiencies for TE and TM polarizations</td>
</tr>
<tr>
<td>3.A7</td>
<td>Reconstruction of the hyperspectral data-cube</td>
</tr>
<tr>
<td>3.A8</td>
<td>Design and fabrication of the hyperspectral test sample</td>
</tr>
<tr>
<td>3.A9</td>
<td>Object reconstruction at different wavelengths by deconvolution</td>
</tr>
<tr>
<td>4.1</td>
<td>Unit cell transmission simulation at different wavelengths</td>
</tr>
<tr>
<td>4.2</td>
<td>Focusing performance of a spherical lens</td>
</tr>
<tr>
<td>4.3</td>
<td>Focusing performance of an optimized lens</td>
</tr>
<tr>
<td>4.4</td>
<td>Simulation of a vortex cornagraph</td>
</tr>
<tr>
<td>4.5</td>
<td>Unit cell, design curves and images of the fabricated vortex</td>
</tr>
</tbody>
</table>
ABBREVIATIONS

2D  two dimensional.
DOE  diffractive optical element.
NA  numerical aperture.
HCA  high-contrast transmit/reflect array.
HCTA  high-contrast transmittarray.
DoFP-PC  division of focal plane polarization camera.
SEM  scanning electron microscope.
PGI  phase gradient image.
DIC  differential interference contrast.
FOV  field of view.
MEMS  microelectromechanical systems.
PSF  point spread function.
α-Si  amorphous silicon.
EBL  electron beam lithography.
OSA  optical spectrum analyzer.
MS  metasurface spectrometer.
DBR  distributed Bragg reflector.
RCWA  rigorous coupled-wave analysis.
PECVD  plasma-enhanced chemical vapor deposition.
Al₂O₃  aluminum oxide.
FWHM  full width at half maximum.
3D  three dimensional.
VC  vortex coronagraph.
EIT  Electromagnetically induced transparency.
DUV  deep ultraviolet.
SLM  spatial light modulator.
Chapter 1

INTRODUCTION

Recognized as thin, flat optical elements with unprecedented functionalities and performances, optical metasurfaces have gained a great deal of interest during the past decade [1]. In this chapter, after a short discussion on the history of optical metasurfaces, their operation principle in contrast to other refractive and/or diffracting optical elements will be investigated. Then, we briefly describe state-of-the-art classes of metasurface devices along with the unique capabilities of each class. Optical systems based on metasurfaces, called metasystems, will be introduced afterwards. Finally a short outline of the thesis will conclude this chapter.

1.1 A history of optical metasurfaces

Optical metasurfaces are 2D arrays of sub-wavelength scatterers which are designed to manipulate different characteristics of light such as its polarization, phase, intensity, or spectral content [2–13]. The subwavelength nano-resonators, called meta-atoms, capture and re-radiate the incident light and depending on how these meta-atom are designed, the scattered light might have different characteristics compared to the incoming incident light. For example, the meta-atoms might impose different phases, polarization states, angles, intensities or spectral contents on the incoming wave. In other words, by precise design of these meta-atoms and placing them in arrays of different forms, the characteristics of the light interacting with the metasurface can be engineered.

Various conventional optical functionalities such as focusing in lenses, diffraction in gratings, reflection in mirrors, holography in holograms, phase retarding in wave-plates, polarization filtering in polarizers, and spectral filters have been realized using optical metasurfaces so far [14–24]. Furthermore, metasurfaces are able to provide multiple functionalities altogether in one surface, which is conventionally only achieved by using a combination of optical components [25]; or entire new functionalities can be assumed for them [26].

Here we review the development of optical metasurfaces with the focus on dielectric metasurfaces for controlling the polarization and wavefront of the light. Optical metasurfaces have a conceptual similarity to the reflect- and/or transmit-array
antennas, which have been studied for a long time in the antenna and microwave community [27]. In particular, early demonstration of the metallic meta-atoms [28] and the use of geometric (or Pancharatnam-Berry) phase for controlling circular polarized light have been known for decades [29, 30]. In addition, many of the metasurface designs and models are extracted from those used before in diffractive optical elements (DOEs) [31–35]. As a result, many of the techniques, theories, and models developed for or used by DOEs [35–40] can be directly applied and/or transformed to be used in the design of optical metasurfaces. At the same time, optical metasurfaces outperform their DOEs counterparts in various regimes including large numerical apertures (NAs).

Optical metasurfaces are typically fabricated in planar forms and are potentially low-cost for mass-production using standard fabrication processes. By the advance of technology in nano-fabrication and wider accessibility of its tools, an increased interest in the research of plasmonic metasurface structures [28, 41–44] as well as high contrast dielectric metasurfaces [15, 45, 46] happened. Some of the early works in this topics are presented in Fig. 1.1.

In recent years, a transition from metallic/plasmonic metasurfaces to dielectric metasurfaces has occurred due to their superior efficiency and lower loss. In most of the earlier plasmonic metasurfaces as shown in Fig. 1.1a the light was manipulated using a single metallic layer by resonance phase, geometric phase or the combination of the two [2, 16, 19, 43, 54–57]. However, the metallic loss and single layer fundamental limitation in the transmission mode [58–60] restrict their performance. The idea of wave-front shaping using the dielectric geometric phase in nano-beam half waveplates [47, 61] has been also investigated as shown in Fig. 1.1b. However operating in single polarization regime and the strong coupling between adjacent elements are the confining factors for achieving high efficiently and large deflection angles. Metallic Huygens’ metasurfaces were introduced to overcome the existing limitations of plasmonic metasurfaces [48, 62, 63], which allow simultaneous excitation of electric and magnetic dipole moments as shown in Fig. 1.1c. Yet, the material loss and complex fabrication are still the prohibiting factors for application of metallic Huygens’ metasurfaces in the optical regime. All dielectric Huygens’ metasurfaces were also introduced [49, 64–69] which allow for two longitudinal resonance modes of electric and magnetic dipole moments with equal frequency as shown in Fig. 1.1d. Still, there were additional challenges related to the practical application and implementation of all-dielectric
Figure 1.1: A short history of optical metasurfaces. (a) Scanning electron micrograph of a plasmonic metasurface beam deflector. Inset shows the interference of vortex beam and a co-propagating Gaussian beam, adapted from [28]. (b) A geometric phase axicon with dielectric micro-bars along with the transversal distribution of Bessel beam in the inset, adapted from [47]. (c) Efficient light bending with metamaterial Huygens’ surfaces, adapted from [48]. (d) Scanning electron micrograph of a dielectric Huygens’ beam deflector. Simulated field intensities showing the electric and magnetic dipole are plotted on the right, adapted from [49]. (e) Scanning electron micrograph of a portion of a high-contrast transmitarray lens, adapted from [50]. (f) Scanning electron micrograph of two high contrast grating lenses, adapted from [51]. (g) Scanning electron micrograph of blazed binary diffractive elements, adapted from [52, 53] from left to right.

Huygens’ metasurfaces. As an example, achieving full $2\pi$ phase coverage and high transmission requires variation of all resonator dimensions including their height which is challenging by planar fabrication technology, and at the same time the mutual coupling between elements deteriorates the performance of the device [49], especially for large deflection angles and fast-varying phase profile. For that reason, the idea of using high-contrast transmit/reflect arrays (HCAs) structures as shown in Fig. 1.1e and f which uses thicker (about $0.5\lambda$ to $\lambda$) high-index layers for patterning the meta-atoms was explored [14, 17, 25, 46, 51, 70–76]. The HCAs structures resemble those blazed binary optical elements [52, 53, 77, 78] as shown in Fig. 1.1g. However they present better performance compared to other classes of metasurfaces.
for wavefront manipulation.

It worth noting that the applications of optical metasurfaces go well beyond spatial wavefront manipulation, incorporating thin light absorbers [79–93], optical filters [20, 22–24, 94–102], nonlinear [103–110], and anapole metasurfaces [111, 112]. In the following section we will discuss state of the art optical HCA metasurfaces for applications mostly focusing on wavefront manipulation.

1.2 State of the art optical metasurface devices

HCAs are practically advantageous compared to other classes of optical metasurfaces. The operating principles of HCAs are based on the confinement of light (electromagnetic energy) inside high-refractive index dielectric nano-scatterers [Fig. 1.2a] surrounded by lower-index media [14, 25, 46, 52, 53, 70, 73–75] and sampling the phase at each nano-scatterer. To achieve the full $2\pi$ phase coverage the cross-section area of nano-scatterers needs to be different. By placing the nano-scatterers in a periodic 2D lattice, the Nuyquist sampling criterion [50] needs to be satisfied to ensure the excitation of the main diffraction mode only. Figures 1.2c and d shows two different types of lattices, square and triangular respectively. [113, 114]. For polarization independence performance the nano-scatterers need to have symmetric cross-sections, in the form of circles [Fig. 1.2a], squares [Fig. 1.2c], hexagons, etc.

In recent years, many demonstrations of high-efficiency high-NA lenses using the HCA have been reported [14, 51, 72, 75, 118, 119]. Figure 1.2a, shows one of the early demonstrations where lenses with measured efficiencies from 82% to 42% depending on the NA (ranging from 0.5 to above 0.95) were demonstrated, while the focal spots are diffraction-limited in each case. As mentioned before, the confinement of light in HCAs would result in negligible coupling between adjacent nano-posts even above the structural cut-off [53]. Furthermore, in contrast to Huygens’ metasurface [49, 64, 68, 120–122], the transmission/reflection behaviour of each nano-post would be independent of its neighboring meta-atom. Hence, the design process for each of the meta-atoms in a phase surface can be performed independent of its adjacent meta-atoms. This assumption will not hold true for high NA large deflection angle metasurfaces, where the efficiency drops at such large incident angles. Some efforts have been made to improve the efficiency at large deflection angles [123–125]. The maximum focusing efficiency for a lens with NA of $\sim$0.8 was measured to be around 75% [124].

Another important property of HCTA-based metasurfaces is their capability to control
Figure 1.2: State of the art optical metasurfaces developed in Faraon group. 
(a) High-NA lenses with high efficiencies [14]. The lenses have NAs from ~0.5 to ~0.97, and measured focusing efficiencies of 82% to 42%. Scale bar: 1μm. 
(b) A polarimetry scheme using rectangular cross-section meta-atoms [115] (left). Required phase profiles for acting as both polarization beam splitter and focusing element at two orthogonal polarizations (top right). Scanning electron micrograph and the three dimensional illustration of superpixel focusing different polarizations to different spots (bottom right). (c) Controlling the chromatic dispersion with metasurfaces [113]. SEM images (right) and reflection phase (center) of reflective squared cross-section nanoposts. (d) Multi-wavelength metasurfaces showing the unit cell (left) along with its optical micrscope and SEM images (right) [116]. 
(e) Metasurfaces with phase and angle control [26]. Schematic of a U-shaped cross-section unit cell (right), the simplified measurement setup (left) and the SEM images of the device (center) 
(f) Conformal optics with optical metasurfaces [117]. A schematic illustration of a dielectric metasurface layer conformed to the surface of a transparent object with arbitrary geometry (center). Optical image of a conformal metasurface mounted on a glass cylinder (left), and overview of the fabrication process (right). (g) Tunable elastic metasurface lenses [50]. Phase profile of a metasurface lens before and after stretching (bottom left) along with the scanning electron micrograph of nano-posts (right). Measured and analytically predicted focal distances for different stretch ratios (center).
light independently based on its various degrees of freedom like polarization \([25, 115, 126, 127]\), wavelength \([113, 114, 116, 128]\), and incident angle \([26, 129]\). Figure 1.2 highlights some of the achieved functionalities of HCTAs based on different degrees of freedom. A few of them will be described in the following:

One of the important properties of HCTAs is the ability to control polarization and phase simultaneously. Polarimetric imaging by measuring the polarization state of light over a scene of interest, provides important information about the molecular and material composition of the scene. A scheme of full Stokes polarimetric imaging by means of a HCA metasurface is shown in Figure 1.2b \([115]\). The scheme called division of focal plane polarization cameras (DoFP-PCs), makes use of the independent phase control for orthogonal polarizations of light.

HCTAs can also be used to alter chromatic diffractive dispersion, a fundamental interference effect dictated by geometry. As shown in Fig. 1.2c by using a reflective metasurface composed of dielectric nano-posts, providing simultaneous control over the phase and its wavelength derivative, diffractive gratings and focusing mirrors with positive, zero, and negative dispersion are demonstrated \([113]\). In addition to the correction for diffractive dispersion, lenses based on HCTAs are designed to operate in multiple wavelengths \([116]\). In this case, the lens behavior at distinct wavelengths is corrected through complete and independent phase coverage at those wavelengths. Figure 1.2d demonstrates the unit cell composed of four scattering elements which provides control on scattering phase at two wavelength of 915 nm and 1150 nm. The optical and scanning electron microscope images of the lens and nano-posts are shown in Fig. 1.2d.

In most of the thin DOEs, the angular response is highly correlated. This fundamental property can be altered by angle-multiplexed metasurfaces, presented in Fig. 1.2e, composed of reflective HCAs of dielectric U-shaped meta-atoms, whose response under illumination from different angles can be independently controlled \([26]\). As shown in Fig. 1.2e by controlling the reflection phases at each incident angle, two distinct hologram images are created under two different incident angles of 0° and 30°.

In most optical components, the optical functionalities and the physical geometries are generally highly correlated. Conformal metasurfaces based on HCTAs can be designed to decouple these two properties \([117]\). Figure 1.2f shows that by conforming the metasurface to the outermost surface of an object, the transmitted wavefront can be shaped in a desired form. As proof of concept, a cylindrical lens
covered with HCTA metasurfaces is transformed to function as an aspherical lens focusing light to a point rather than a line.

Elastic highly tunable optical metasurfaces based on HCTAs encapsulated in a thin transparent polymer are also realized \[50\]. Figure 1.2g shows that the focal distance can be tuned by more than 130\%, (from 600 μm to 1400 μm) while maintaining high efficiency and diffraction limited focus.

The extensive ability of HCAs in realizing high NA lenses with high efficiencies as well as their success in achieving the above-discussed functionalities, have been the reason for the great attention toward them in the past few years. At the same time, their compactness and planar form factor as well as the compatibility with micro- and nano-fabriation techniques make HCAs a suitable platform for realizing compact optical systems, either in the form of vertically cascaded metasurfaces \[130–132\], folded structures \[133, 134\], or other forms. In the next section, we will briefly review HCA optical systems composed of more than one optical metasurfaces.

### 1.3 Optical meta-systems

So far, the focus of this chapter has been on single-metasurface components. However, the true potential of metasurfaces can only be reached when multiples metasurfaces are co-designed and integrated to achieve an optical system called meta-system. In this section, this capability will be demonstrated through several examples providing functionalities not feasible with single metasurfaces due to limited degrees of freedom. The discussed examples in this section show vertical cascaded integration of metasurfaces for applications such as correcting the geometric aberration of imaging lenses \[130\], flat optical retro-reflector \[131\], multi-spectral chromatic optics \[135\], etc.

Vertically cascaded metasurfaces can perform sophisticated image corrections which can then be directly integrated on top of image sensors. A miniature monolithic metasurface doublet lens corrected for monochromatic aberrations is demonstrated in Fig. 1.3a \[130\]. The doublet lens which acts as a fisheye photographic objective, has a small $f$-number of 0.9, an angle-of-view larger than $60^\circ \times 60^\circ$, and operates at 850 nm wavelength with 70\% focusing efficiency. The camera exhibits nearly diffraction-limited performance, shown in Fig. 1.3, which indicates the potential of the technology for development of optical microscopy, photography and/or computer vision systems.

The systems of vertically cascaded metasurfaces, each performing a pre-defined
Figure 1.3: Optical meta-systems. (a) Schematic illustration of a monolithic metasurface doublet lens (top) [130]. Simulated focal plane intensity for different incident angles (bottom) (b) Monolithic planar retroreflector made of two metasurfaces. Optical image of an array of retroreflectors along with an SEM image of the nano-posts composing the metasurfaces (top). The captured reflectance profile of the retroreflector as the incident angle was varied by rotating the retroreflector (bottom) [131]. (c) Artist’s view of the three-layer metasurface lens (left) [135]. Dark-field images of the single-layer metasurfaces which are designed to focus, green or blue to the focal distance of 1 mm (right). (d) Schematic illustration of the tunable doublet lens [132]. Schematic illustration of the proposed tunable lens, comprised of a stationary lens on a substrate, and a moving lens on a membrane (left). Intensity distributions in the focal plane of the doublet lens at different actuation voltages (top right). Measured front focal length versus the applied DC voltage (bottom right). (e) Schematic of a QPGM employing two metasurface layers (top). Schematic of the miniaturized optical set-up showing a CMOS image sensor and a linear polarizer attached to it (bottom left) along with PGIs calculated from the three DIC images [136]. (f) Schematic of a metastructure which focuses two different wavelengths to two separate focal points (top). scanning electron microscopes (SEMs) of the cross-sectional view of the 2.5D metalens (bottom left) along with the image plane intensity distributions measured at 780 and 915 nm wavelengths (bottom right) [137].
mathematical transformation [138] have also great potentials for implementing novel high performance components and enabling low-power and low-weight passive optical transmitters [139–141]. As an example, the retro-reflection functionality can be achieved with two vertically stacked planar metasurfaces, the first performing a spatial Fourier transform and its inverse, and the second imparting a spatially varying momentum to the Fourier transform of the incident light [131]. The concept is shown in Fig. 1.3b, where a planar monolithic near-infrared retroreflector, made of two layers of silicon nano-posts is shown. The realized retro-reflection reflects light along its incident direction and has a normal incidence efficiency of 78%, and a large half power FOV of 60°.

Another example of vertically cascaded metasurfaces used to alleviate the chromatic aberrations of individual diffractive elements is presented in Fig. 1.3c, showing a dense vertical stacking of independent metasurfaces, each made of a different material and designed for a different spectral band [135]. Using this approach, a triply red, green and blue achromatic metalens in the visible range is demonstrated.

Furthermore, the integration of optical metasurfaces with the MEMS technology has great potentials for realization of fast tunable and reconfigurable optics [132]. Figure 1.3d shows the concept of a tunable metasurface doublet based on MEMS, with more than 60 diopters (about 4%) change in the optical power upon a 1-μm movement of one metasurface, and a scanning frequency in the order of a few kHz.

Quantitative phase imaging (QPI) of transparent samples which plays an essential role in multiple numerous bio-medical applications, can be realized using a system of optical metasurfaces as shown in Fig. 1.3e [136]. The presented quantitative phase gradient microscope (QPGM) is based on two dielectric metasurface layers having a total volume of ~1-mm³ and captures three DIC images to generate a quantitative phase gradient image in a single shot.

One way to provide additional degrees of freedom and implement efficient multi-functional devices is by using stacked layers of interacting meta surface layers such as 2.5D metastructures. As an example, a 2.5D structure designed using the adjoint optimization technique is demonstrated in Fig. 1.3f [137]. In this technique, instead of designing the unit cells individually, the structure is considered as a whole, which accounts for inter-post as well as inter-layer coupling. As shown in Fig. 1.3f, the bilayer meta-structure is a bi-layer lens designed to focus two incident wavelengths of 780 nm and 915 nm to two points 60 μm away from the structure and ~20 μm apart from each other.
All of the meta-systems presented in this section obey the vertically cascaded topology. In the next chapter we introduce another way for implementing optical meta-systems using folded metasurface optics.

1.4 Thesis outline
Chapter 2 of this dissertation is focused on proposing the folded metasurface optics platform as a compact and robust configuration for realizing advanced optical systems. The proposed platform, significantly reduces the size of optical systems. To demonstrate the practicality of the presented platform, a compact high-resolution folded metasurface spectrometer operating in near-IR is showcased.

Chapter 3 discusses multiple ways for implementing miniaturized hyper-spectral imagers, among which a scheme based on the folded metasurface optics platform was designed and implemented. In addition, a computational compact hyperspectral imaging scheme using a metasurface phase mask is introduced and two reconstruction techniques for obtaining the hyperspectral data-cube are investigated.

In Chapter 4, other potentials of optical metasurfaces, including the optimization techniques to design highly-efficient devices are investigated. As a not-yet-fully-explored application, the design and implementation of optical metasurface vortex cornagraphs have been explored.

We conclude this dissertation by discussing the potentials and challenges facing optical metasurface devices and systems, and presenting a vision towards the prospective future for the field in Chapter 5.
The material in this chapter was in part presented in [133, 142].

2.1 Compact Folded metasurface spectrometer

An optical design space that can highly benefit from the recent developments in metasurfaces is the folded optics architecture where light is confined between reflective surfaces, and the wavefront is controlled at the reflective interfaces. Here, we introduce the concept of folded metasurface optics by demonstrating a compact spectrometer made from a 1-mm-thick glass slab with a volume of 7 cubic millimeters. The spectrometer has a resolution of \( \sim 1.2 \) nm, resolving more than 80 spectral points from 760 to 860 nm. The device is composed of three reflective dielectric metasurfaces, all fabricated in a single lithographic step on one side of a substrate, which simultaneously acts as the propagation space for light. The folded meta-system design can be applied to many optical systems, such as optical signal processors, interferometers, hyperspectral imagers and computational optical systems, significantly reducing their sizes and increasing their mechanical robustness and potential for integration.

Introduction

Optical spectrometry is a key technique in various areas of science and technology with a wide range of applications [143, 144]. This has resulted in a large demand for spectrometers and/or spectrum analyzers with different properties (e.g., operation bandwidth, resolution, size, etc.) required for different applications [145–147]. Conventional optical spectrometers are composed of a dispersive element, such as a prism or a diffraction gating, that deflects different wavelengths of light by different angles, followed by focusing elements that focus light incoming from different angles to different points (or lines). As schematically shown in Fig. 2.1a, the intensity at different wavelengths can then be measured using an array of detectors. Diffraction gratings have typically larger dispersive powers than transparent materials, and therefore diffractive spectrometers generally have better resolutions [143]. The combination of several free space optical elements (the grating, focusing mirrors, etc.) and the free space propagation volume result in bulky spectrometers. In recent years,
there has been an increased interest in high-performance compact spectrometers that can be easily integrated into consumer electronics for various medical and technological applications such as medical diagnosis, material characterization, quality control, etc. [148, 149]. As a result, various schemes and structures have been investigated for realization of such spectrometers [149–158]. One class of miniaturized spectrometers integrate a series of band-pass filters with different center wavelengths on an array of photodetectors [99, 150]. Although these devices are compact and compatible with standard micro-fabrication techniques, they have resolutions limited by achievable filter quality factors, and low sensitivities caused by the filtering operation that rejects a large portion of the input power. Spectrometers based on planar on-chip integrated photonics provide another solution with high spectral resolution [149, 151–155]. However, the loss associated with on-chip coupling of the input light and the reduced throughput because of the single-mode operation [159] are still major challenges for widespread adoption in many applications.

Another type of compact spectrometers are conceptually similar to the conventional table-top spectrometers, however, they use micro-optical elements to reduce size and mass [156, 157]. Due to the inferior quality and limited control achievable by micro-optical elements as well as the shorter optical path lengths, these devices usually have lower spectral resolutions. Higher resolution has been achieved by using aberration-correcting planar gratings [158], however an external spherical mirror makes the device bulky.

A key feature of metasurfaces is their compatibility with micro and nano-fabrication techniques, which allows for integration of multiple metasurfaces for realizing complex optical metasystems [130, 131]. Such metasystems allow for significantly improving optical properties of metasurfaces through aberration correction (such as lenses with diffraction-limited operation over wide FOV [130]), or functionalities fundamentally unachievable with local single-layer metasurfaces such as retroreflection [131].

**Concept and design**

Taking a different approach to device integration, here we introduce folded optical metasystems where multiple metasurfaces are integrated on a single substrate that is also playing the role of propagation space for light [Fig. 2.1b]. Using this platform, we experimentally demonstrate a compact folded optics device for spectroscopy with a 1-mm thickness (∼7-mm³ volume) that provides a ∼1.2-nm resolution over
Figure 2.1: Schematics of a conventional and a folded metasurface spectrometer. (a) Schematic illustration of a typical diffractive spectrometer. The main components are comprised of the fore-optics section, diffraction grating, focusing lenses and detector array. (b) The proposed scheme for a folded compact spectrometer. All the dispersive and focusing optics can be implemented as reflective metasurfaces on the two sides of a single transparent substrate. Mirrors on both sides confine and direct light to propagate inside the substrate, and the detector can be directly the output aperture of the device. If required, transmissive metasurfaces can also be added to the input and output apertures to perform optical functions. Although the schematic here includes metasurfaces on both sides to show the general case, the actual devices demonstrated here are designed to have metasurfaces only on one side to simplify their fabrication.

A 100-nm bandwidth (more than 80 points over a ~12% bandwidth) in the near infrared. As schematically shown in Fig. 2.1b, multiple reflective metasurfaces can be designed and fabricated on the same transparent substrate to disperse and focus light to different points on a plane parallel to the substrate. The folded structure resembles the integrated optical module presented in [160] to provide better mechanical alignment of optical components. Yet, to the best of our knowledge, this is the first demonstration of an optical metasystem comprising more than two metasurfaces that implements a sophisticated optical functionality like spectrometry. Furthermore, the presented configuration allows for the integration of the detector array on top of the folded spectrometer, resulting in a compact monolithic device. We should note here that it was recently demonstrated that an off-axis metasurface lens
(i.e., a lens with an integrated blazed grating phase profile [161, 162]) can disperse and focus different wavelengths to different points. However, there are fundamental and practical limitations for such elements that significantly limits their application as a spectrometer (which is the reason why other types of diffractive optical elements, such as holographic optical elements and kinoforms, that can essentially perform the same function have not been used for this application before). Fundamentally, the chromatic dispersion [113, 114, 163–165] and angular response correlation of diffractive optical elements and metasurfaces [26, 166] limit the bandwidth and angular dispersion range where the device can provide tight aberration-free focusing. This in turn limits the achievable resolution and bandwidth of the device. Moreover, the chromatic dispersion results in a focal plane almost perpendicular to the metasurface, which will then require the photodetector array to be placed almost normal to the metasurface plane [161, 162, 167]. In addition to the distance for the propagation of dispersed light, this normal placement undermines the compactness of the device.

Figure 2.2a shows the ray tracing simulations of the designed spectrometer. The device consists of three metasurfaces, all patterned on one side of a 1-mm-thick fused silica substrate. The first metasurface is a periodic blazed grating with a period of 1 μm that disperses different wavelengths of a collimated input light to different angles, centered around 33.9° at 810 nm. The second and third metasurfaces focus light coming from different angles (corresponding to various input wavelengths) to different points on the focal plane. We have recently demonstrated a metasurface doublet capable of correcting monochromatic aberrations to achieve near-diffraction-limited focusing over a wide FOV [130]. The second and third metasurfaces here essentially work similar to the mentioned doublet, with the difference of working off axis and being designed in a folded configuration, such that the focal plane for our desired bandwidth is parallel to the substrate. To simplify the device characterization, the focal plane was designed to be located ~200 μm outside the substrate. The asymmetric design of the focusing metasurfaces in an off-axis doublet configuration allows for the focal plane to be parallel to the substrate. This makes the integration of the spectrometer and the detector array much simpler, results in a more compact and mechanically robust device, and allows for direct integration into consumer electronic products like smartphones. The optimized phase profiles for the two surfaces are shown in Fig. 2.2a, right (see Table 2.A1 for the analytical expression of the phases). Simulated spot diagrams of the spectrometer are plotted in Fig. 2.2b for three wavelengths at the center and the two ends of the bandwidth showing negligible
Figure 2.2: Ray-optics design and simulation results of the folded spectrometer. 
(a) Ray tracing simulation results of the folded spectrometer, shown at three wavelengths in the center and two ends of the band. The system consists of a blazed grating that disperses light to different angles, followed by two metasurfaces optimized to focus light for various angles (corresponding to different input wavelengths). The grating has a period of 1 μm, and the optimized phase profiles for the two metasurfaces are shown on the right. (b) Simulated spot diagrams for three wavelengths: center and the two ends of the band. The scale bars are 5 μm. (c) Spectral resolution of the spectrometer, which is calculated from simulated Airy disk radii and the lateral displacement of the focus with wavelength. (d) Simulated intensity distribution for two wavelengths separated by 1.1 nm around three different center wavelengths of 760 nm, 810 nm, and 860 nm. The intensity distributions show that wavelengths separated by 1.1 nm are theoretically resolvable. The scale bars are 20 μm.

geometric aberrations. The spot diagrams are plotted only at three wavelengths, but the small effect of optical aberrations was confirmed for all wavelengths in the 760–860 nm bandwidth. As a result, the spectral resolution of the device can be calculated using the diffraction-limited Airy radius and the lateral displacement of the focus by changing the wavelength. The calculated resolution is plotted in Fig. 2.2c, showing a theoretical value of better than 1.1 nm across the band. Point spread functions (PSFs) calculated for input beams containing two wavelengths 1.1 nm apart, and centered at 760 nm, 810 nm, and 860 nm are plotted in Fig. 2.2d, showing two resolvable peaks.

To implement the reflective metasurfaces, we used a structure similar to the reflective elements in [131]. Each of the meta-atoms, shown schematically in Fig. 2.3a, consists of an α-Si nano-post with a rectangular cross section, capped by a ~2-
**Figure 2.3: Metasurface structure and design graphs.** (a) Schematics of the reflective rectangular meta-atom. The meta-atom consists of α-Si nano-posts on a fused silica substrate, capped by a layer of SU-8 polymer and backed by a gold mirror. The nano-post is 395 nm tall and the lattice constant is 250 nm for the blazed grating and 246 nm for the focusing metasurfaces. Schematics of the simulated structure and conditions are shown on the bottom. (b) and (c) Simulated reflection phase plotted for TE and TM polarizations. The black curve highlights the path through the $D_x-D_y$ plane that results in equal phases for the two polarizations. Nano-posts on this path were used to design the two focusing metasurface elements to make them insensitive to the input polarization.

μm-thick SU-8 layer and backed by a gold mirror. The post height and lattice constant were chosen to be 395 nm and 246 nm, respectively, to achieve full $2\pi$ phase coverage while minimizing variation of the reflection phase derivative across the band [Fig. 2.A1]. Minimizing the phase derivative variation will mitigate the reduction of device efficiency over the bandwidth of interest [113] by decreasing the wavelength dependence of the phase profiles [Fig. 2.A2]. In addition, since the two focusing metasurfaces are working under an oblique illumination ($\theta \sim 33.9^\circ$), the nano-posts were chosen to have a rectangular cross-section to minimize the difference in reflection amplitude and phase for the TE and TM polarizations (for the oblique incidence angle of 33.9° at 810 nm). Reflection coefficients are found through simulating a uniform array of nano-posts under oblique illumination ($\theta \sim 33.9^\circ$) with TE and TM polarized light [Fig. 2.3a, right]. The simulated reflection phase as a function of the nano-posts side lengths are shown in Fig. 2.3b and 2.3c for TE and TM polarizations. The black triangles highlight the path through the $D_x-D_y$ plane along which the reflection phase for the TE and TM polarizations is almost equal. In addition, as shown in Fig. 2.A2, having almost the same reflection phases for the TE and TM polarizations holds true for the whole desired 760–860 nm bandwidth. The nanopost dimensions calculated from this path were used to implement the two focusing metasurfaces.
The blazed grating has a periodic phase profile (with a period of ~1 μm) that deflects normally incident light to a large angle inside the substrate. With a proper choice of the lattice constant (250 nm, in our case), its structure can also be periodic. This different structure and operation require a different design approach. The periodicity of the grating allows for its efficient full-wave simulation which can be used to optimize its operation over the bandwidth of interest. A starting point for the optimization was chosen using the recently developed high-NA lens design method [124], and the structure was then optimized using the particle swarm optimization algorithm to simultaneously maximize deflection efficiency at a few wavelengths in the band for both polarizations (see Appendix 2.2 and Fig. 2.A3 for details).

Device fabrication

The device was fabricated using conventional micro- and nano-fabrication techniques. First, a 395-nm-thick layer of α-Si was deposited on a 1-mm-thick fused silica substrate. All metasurfaces were then patterned using EBL in a single step, followed by a pattern inversion through the lift-off and dry etching processes. The metasurfaces were capped by a ~2 μm-thick SU-8 layer, and a 100-nm-thick gold layer was deposited as the reflector. A second reflective gold layer was deposited on the second side of the substrate. Both the input and output apertures (with diameters of ~790 μm and ~978 μm, respectively) were defined using photolithography and lift-off. An optical microscope image of the three metasurfaces, along with a scanning electron micrograph of a part of the fabricated device are shown in Fig. 2.4a.

Experimental and characterization results

To experimentally characterize the spectrometer, a normally incident collimated beam from a tunable continuous wave laser was shinned on the input aperture of the device. A custom-built microscope was used to image the focal plane of the spectrometer, ~200 μm outside its output aperture (see Appendix 2.2 and Fig. 2.A5 for details of the measurement setup). The input wavelength was tuned from 760 to 860 nm in steps of 10 nm, and the resulting intensity distributions were imaged using the microscope. The resulting one-dimensional intensity profiles are plotted in Figs. 2.4b and 2.4c for TE and TM polarizations. The intensity profiles were measured over the whole 1.2-mm length of the y-direction in the focal plane (as shown in Fig. 2.4, inset) at each wavelength. The background intensity is beyond visibility in the linear scale profiles plotted here for all wavelengths (see Figs. 2.A6 and 2.A7 for two-dimensional and logarithmic-scale plots of the intensity distribution, respectively). Figures 2.4d and 2.4e show the measured intensity profiles for three
Figure 2.4: Experimental characterization results. (a) An optical microscope image of the fabricated device and metasurfaces before deposition of the second gold layer. Inset shows a scanning electron micrograph of a portion of one of the two focusing metasurfaces (scale bars from left to right: 10 mm, 500 μm, and 1 μm). (b) and (c) One dimensional focal spot profiles measured for several wavelengths in the bandwidth along the y-direction (as indicated in the inset) for TE and TM polarizations. The wavelengths start at 760 nm (blue curve) and increase at 10-nm steps up to 860 nm (red curve). (d) and (e) Measured intensity distributions for two input wavelengths that are 1.25 nm apart for TE and TM polarizations. The measurements were carried out at the center and the two ends of the wavelength range for both polarizations. The insets show the corresponding 2-dimensional intensity profiles, demonstrating two resolvable peaks (scale bars: 10 μm). (f) Calculated and measured absolute focusing efficiencies of the spectrometer for TE and TM polarizations. Both polarizations have average measured efficiencies of ~25%.
sets of close wavelengths, separated by 1.25 nm. The insets show the corresponding
two-dimensional intensity distribution profiles. For all three wavelengths, and for
both polarizations the two peaks are resolvable. The experimentally obtained spectral
resolution is plotted in Fig. 2.A8 versus the wavelength. The average resolution for
both polarizations is \( \sim 1.2 \) nm, which is slightly worse than the theoretically predicted
value (\( \sim 1.1 \) nm). We attribute the difference mostly to practical imperfections such as
the substrate having an actual thickness different from the design value and thickness
variation. In addition the metasurface phases are slightly different from the designed
profiles due to fabrication imperfections. The angular sensitivity/tolerance of the
device was also measured with respect to polar and azimuthal angle deviations from
0 incidence angle, in the \( x-z \) and \( y-z \) planes, using the setup shown in Fig. 2.A9c. In
the \( y-z \) plane the maximum tilt angle to maintain the same 1.25 nm resolution is
\( \pm 0.15^\circ \), while in the \( x-z \) plane the device has a \( \pm 1^\circ \) degree acceptance angle. The
measurement results in Fig. 2.A9 match well with the predictions from ray-tracing
simulations.

The measured and calculated focusing efficiencies are plotted in Fig. 2.4f. The
focusing efficiency, defined as the power passing through a \( \sim 30 \) \( \mu \)m diameter
pinhole around the focus divided by the total power hitting the input aperture, was
measured using the setup shown in Fig. 2.A5. For both polarizations, the average
measured efficiency is about 25%. As seen from the measured efficiency curves, the
optimization of the blazed grating efficiency versus wavelength and the choice of
the design parameters to minimize variations in the phase-dispersion for the doublet
metasurface lens, have resulted in a smooth measured efficiency. An estimate for
the expected efficiency (shown as simulated efficiency in Fig. 2.4f) is calculated by
multiplying the deflection efficiency of the grating, the efficiency of seven reflections
off the gold mirrors, the input and output aperture transmission efficiencies, and
the average reflectivities of the uniform nano-post arrays (as an estimate for the
two focusing metasurface efficiencies). It is worth noting that considering only the
reflection losses at the interfaces (nine reflective ones, and two transmissive ones)
reduces the efficiency to about 48%, showing a close to 50% efficiency for the three
metasurfaces combined. We attribute the remaining difference between the measured
and estimated values to fabrication imperfections (e.g., higher loss for the actual
gold mirrors, and imperfect fabrication of the metasurfaces), the lower efficiency of
metasurfaces compared to the average reflectivity of uniform arrays, and to the minor
difference from the designed value of the metasurface phase profiles at wavelengths
other than the center frequency.
Figure 2.5: Sample spectrum measurement results. (a) Spectrum of a wideband source (a super-continuum laser with an 840-nm short-pass filter) measured by a commercial OSA and the MS. This measurement was used to extract the calibration curve. (b) Spectrum of a 10-nm band-pass filter centered at 800 nm measured by the MS, compared to the spectrum acquired from the filter datasheet. (c) Transmission spectrum of a Nd : YVO₄ crystal sample measured both with a commercial OSA and the MS. (d) The optical depth of the sample extracted from the spectrum measurements both with the OSA and the MS.

Spectrum Measurement

To demonstrate that the metasurface spectrometer actually has the ability to measure dense optical spectra, we use it to measure the transmission spectra of two different samples. First, we measured the spectrum of a wideband source (a super-continuum laser source, filtered with an 840-nm short-pass filter), both with the metasurface spectrometer (MS) and a commercial optical spectrum analyzer (OSA). By dividing the spectra measured by the two devices, we extract the required calibration curve that accounts for the variation of the metasurface spectrometer as well as the non-uniformities in the responsivity of the optical setup used to image the focal plane (i.e., the objective lens and the camera, as well as the optical fiber used to couple the signal to the OSA). The measured spectra and the extracted calibration curve are plotted in Fig. 2.5a. Next, we used this calibration curve to measure the transmission spectrum of a band-pass filter with a nominal 10-nm full width at half maximum bandwidth and centered at 800 nm. The measured spectrum along with the transmission spectrum obtained from the filter datasheet are plotted in Fig. 2.5b, showing a good agreement.
Finally, we used the metasurface spectrometer to measure the optical depth of a Nd:YVO$_4$ crystal sample. The spectrum measured with the metasurface spectrometer (after calibration) is compared with the transmission spectrum of the same sample measured with the OSA in Fig. 2.5c. Dividing the spectrum without and with the sample, we have extracted the optical depth of the sample which is plotted in Fig. 2.5d. A good agreement is observed between the two measurement results. It is worth noting that the Nd:YVO$_4$ crystal sample was cut though the z-plane, resulting in an equal absorption spectrum for the two polarizations. Therefore, we can assume that all spectral measurements were done with the same state of input polarization. This justifies the use of only one calibration curve for all the measurements.

**Discussion**

The measured efficiency of the spectrometer demonstrated here is about 25%. This value can be significantly increased to about 70% by using mirrors with higher reflectivity (e.g., DBRs or high contrast grating mirrors [168, 169]), and anti-reflection coatings on the input and output apertures. In addition, more advanced optimization techniques [125] could be exploited to optimize the diffraction grating to achieve high efficiency and polarization insensitivity. Implementing these changes and optimizing the fabrication process, one can expect to achieve efficiencies exceeding 70% for the spectrometer.

The metasurface spectrometers are fabricated in a batch process, and therefore many of them can be fabricated on the same chip, even covering multiple operation bandwidths. This can drastically reduce the price of these devices, allowing for their integration into various types of systems for different applications. In addition, the demonstrated structure is compatible with many of the techniques developed for the design of multi-wavelength metasurfaces [114, 116], and therefore one might be able to combine different optical bandwidths into the same device (e.g., using a grating that deflects to the right at one bandwidth, and to the left at the other), resulting in compact devices with enhanced functionalities.

The optical throughput (etendue) is a fundamental property of any optical system, setting an upper limit on the ability of the system to accept light from spatially incoherent sources. It can be estimated as the product of the physical aperture size and the acceptance solid angle of the system. Furthermore, the total etendue of a system is limited by the element with the lowest etendue. To calculate the throughput of the metasurface spectrometer, we have performed simulations and measurements to characterize its acceptance angle. According to the measurement results in
Fig. 2.A9 the acceptance angle of the system is about 2 degrees in the horizontal direction, and 0.3 degrees in the vertical direction. Given this and the input aperture dimensions, the optical throughput of our device is calculated to be $\sim 90 \text{ Sr} \mu\text{m}^2$. For comparison, the etendue of optical systems operating around $\sim 1 \mu\text{m}$ that utilize single-mode input channels (i.e., most optical spectrometers based on integrated optics platforms) is around $\sim 1 \text{ Sr} \mu\text{m}^2$. Furthermore, the demonstrated spectrometer is optimized for maximum sensitivity and not throughput. To show that the achieved throughput here does not denote an upper limit for the etendue of a folded metasurface spectrometer with similar characteristics (i.e., resolution, bandwidth, etc.), we have designed a second device with a throughput of $\sim 13000 \text{ Sr} \mu\text{m}^2$ (see Appendix 2.2 and Fig. 2.A10 for design details and simulation results). Table 2.1 provides a comparison of the optical throughput of several compact spectrometers from recent literature. According to Table 2.1, the spectrometers designed using the folded metasurface platform can collect 2 to 4 orders of magnitude more light compared to on-chip spectrometers that are based on single/few-mode input waveguides, resulting in a much higher sensitivity.

Table 2.1: **Comparison of different spectrometers in terms of throughput (etendue) and their dimensions**

<table>
<thead>
<tr>
<th>Spectrometer</th>
<th>Etendue [Sr$\mu$m$^2$]</th>
<th>Size (dimensions)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[153]</td>
<td>&lt;0.5</td>
<td>50 $\mu$m $\times$ 100 $\mu$m $\times$ thickness</td>
</tr>
<tr>
<td>[170]</td>
<td>$\sim$ 0.8</td>
<td>16 mm $\times$ 7 mm $\times$ 15 $\mu$m</td>
</tr>
<tr>
<td>[157]</td>
<td>8250</td>
<td>20.1 mm $\times$ 12.5 mm $\times$ 10.1 mm</td>
</tr>
<tr>
<td>This work</td>
<td>$\sim$ 90</td>
<td>1 mm $\times$ 1 mm $\times$ 7 mm</td>
</tr>
<tr>
<td>Fig. 2.A10</td>
<td>$\sim$13000</td>
<td>2 mm $\times$ 2.5 mm $\times$ 8 mm</td>
</tr>
</tbody>
</table>

The development of thin and compact optical elements and systems has been a key promise of optical metasurfaces. Although many optical devices have been developed in thin and compact form factors using metasurfaces, significantly reducing the volume of optical systems using metasurfaces has not been previously demonstrated due to the requirement of the free-space propagation in many systems (e.g., imaging systems, spectrometers, etc.). The folded metasystem configuration introduced here can significantly reduce the size of many of these optical systems using the substrate as the propagation space for light. Based on this configuration, we demonstrated a 1-mm-thick spectrometer with a 7-mm$^3$ volume, reduced by a factor of ten compared to the same system implemented in an unfolded scheme (twenty times reduction, if the same system was designed in air). The spectrometer has a resolution of $\sim 1.2 \text{ nm}$ over a 100-nm bandwidth ($> 12\%$) in the near infrared. Using this design,
multiple spectrometers can be fabricated on the same chip and in the same process, significantly reducing the costs and enabling integration of spectrometers covering multiple optical bands into consumer electronics. Moreover, by improving the angular response of the current device one can design a compact hyperspectral imager capable of simultaneous one-dimensional imaging and spectroscopy. In a broader sense, we expect that the proposed platform will also be used for on-chip interferometers, imaging systems, and other devices performing complex transformations of the field.

2.2 Appendix: Additional information and discussion for the metasurface spectrometer

Simulation and design

Ray tracing simulations of the spectrometer were performed using Zemax OpticStudio. In the simulations, metasurfaces were assumed to be phase-only diffractive surfaces. The grating was modeled as a blazed grating with a linear phase along the direction of dispersion (y), and independent of the other direction (x). The phase was chosen to correspond to a period of 1 μm, resulting in deflection angles of 31.6° and 36.35° at 760 nm and 860 nm, respectively. The angles were chosen such that the focused light could be captured by an objective with a NA of 0.95, while maximizing the dispersive power. The second and third surfaces were modeled as a summation of Cartesian coordinate polynomials (Binary 1), $\sum_{n,m} a_{m,n} x^m y^n$, and cylindrical coordinate radially symmetric polynomials (Binary 2) $\Sigma b_{2i} r^{2i}$. The coefficients were optimized to reduce geometric aberrations by minimizing the root mean square geometric spot radii for several input wavelengths covering the bandwidth. The optimized coefficients are given in Table 2.A1. As shown in Fig. 2.2b, all focal spots are optimized and are within the airy disks. This indicates that the designed spectrometer has small geometrical aberrations. The diffraction-limited resolution curve obtained is shown in Fig. 2.2c. The simulations and optimizations were first performed in an unfolded configuration for simplicity. There were several constraints in finding the sizes for input and output apertures. Two opposing factors existed in determination of the 790-μm input aperture diameter. On one hand, a larger input aperture results in a higher throughput and more captured light as well as a higher NA and potentially better resolution. On the other hand, the aperture size for the folded platform cannot be arbitrarily large because different metasurfaces should not overlap. Thus, the 790-μm aperture diameter was chosen in the ray-tracing simulations as the largest size for which metasurface overlap can be avoided and
diffraction-limited focusing can be achieved. The output aperture spatially filters the out of band wavelengths while passing through the bandwidth of interest. Therefore, its size was chosen as the smallest possible aperture that allows for all wavelengths of interest to pass through. Using the ray-tracing simulations, this optimum size was found to be 978 μm.

The RCWA technique [171] was used to obtain reflection phases of the nano-posts. For each specific set of dimensions, a uniform array of the α-Si nano-posts was illuminated with a plane wave at the wavelength of 810 nm under an illumination angle of 33.9° and the reflected amplitudes and phases were extracted for each polarization. To choose the height of the nano-posts, we performed these simulations for nano-posts with square cross-sections and different heights and side lengths [Fig. 2.A1]. The height was then chosen to minimize the variation of the derivative of the phase with respect to wavelength for different side lengths, while providing a full $2\pi$ phase coverage and high reflectivity. Considering the results of Fig. 2.A1b and Fig. 2.A1d, we chose the thickness to be 395 nm. Although this height is slightly less than $\lambda/2$, it is large enough to provide a full $2\pi$ phase coverage as the device operates in reflection mode. The lattice constant was chosen to be 246 nm in order to satisfy the sub-wavelength condition and avoid higher order diffractions, which require $l_c < \lambda/n(1 + \sin(\theta_{\text{max}}))$, where $l_c$ is the lattice constant, $n$ is the refractive index of the substrate, and $\theta_{\text{max}}$ is the maximum deflection angle [50]. We chose $\sin(\theta_{\text{max}}) = 1/n$, since light traveling at larger angles will undergo total internal reflection at the output aperture. To make the two focusing metasurfaces polarization-insensitive, reflection phase and amplitudes were obtained for nano-posts with rectangular cross section under oblique illumination with both TE and TM polarizations [Fig. 2.3]. The design curves were then generated by determining a path in the $D_x$-$D_y$ plane along which TE and TM reflection phases are almost equal.

For designing the blazed diffraction grating, we chose to use the same α-Si thickness of 395 nm (for ease of fabrication). The lattice constant was set to be 250 nm, such that a grating period contains four nano-posts, and the structure becomes fully periodic. This allows for using periodic boundary conditions in the full-wave simulations of the structure, reducing the simulation domain size significantly. The initial values of the post widths were chosen using a recently developed high-NA metasurface design approach [124]. The simulation results for nano-post-width vs reflection-phase and the initial post widths are plotted in Fig. 2.A3a. These values were then fed to a particle swarm optimization algorithm (using an RCWA forward
solver) as a starting point. The algorithm optimizes the deflection efficiency of the grating for both polarizations at 11 wavelengths spanning the bandwidth of interest. The optimization parameters are the side lengths of the rectangular nano-posts, while their thickness and spacing is fixed. Deflection efficiencies of the initial and optimized gratings are plotted in Fig. 2.A3b. The corresponding nano-post widths for both gratings are given in Table 2.A2.

Sample fabrication

A summary of the key steps of the fabrication process is shown in Fig. 2.A4. A 395-nm-thick layer of $\alpha$-Si was deposited on one side of a 1-mm-thick fused silica substrate through a PECVD process at 200°C. The metasurface pattern was then generated in a ~300-nm-thick layer of ZEP-520A positive electron resist (spun for 1 minute at 5000 rpm) using an EBPG5200 EBL system. After development of the resist in a developer (ZED-N50, Zeon Chemicals), a ~70-nm-thick Al$_2$O$_3$ layer was evaporated on the sample in an electron beam evaporator. After lift-off, this layer was used as a hard mask for dry etching the $\alpha$-Si layer in a mixture of SF$_6$ and C$_4$F$_8$ plasma. The alumina layer was then removed in a 1:1 solution of H$_2$O$_2$ and NH$_4$OH. A ~2-$\mu$m-thick layer of SU-8 2002 polymer was spin-coated, hard-baked, and cured on the sample to protect the metasurfaces. The output aperture (which is on the same side as the metasurfaces) was defined using photolithography (AZ-5214E positive resist, MicroChemicals) and lift-off. A ~100-nm-thick gold layer was deposited as the reflective surface. To protect the gold reflector, a second layer of SU-8 2002 was used. To define the input aperture, a ~2-$\mu$m-thick layer of SU-8 2002 polymer was spin-coated and cured on the second side of the wafer to improve adhesion with gold. The input aperture was then defined in a process similar to the output aperture.

Device characterization procedure

The measurement setups used to characterize the spectrometer are schematically shown in Fig. 2.A5. Light from a tunable Ti:sapphire laser (SolsTiS, M-Squared) was coupled to a single mode optical fiber and collimated using a fiber collimator (F240FC-B, Thorlabs). A fiber polarization controller and a free space polarizer (LPVIS100-MP2, Thorlabs) were used to control the input light polarization, and different neutral density filters were used to control the light intensity. The beam illuminated the input aperture of the spectrometer at normal incidence. The focal plane of the spectrometer, located ~200 $\mu$m away from the output aperture, was then imaged using a custom built microscope (objective: 100× UMPiPlanFl, NA=0.95, Olympus; tube lens: AC254-200-C-ML, Thorlabs; camera: CoolSNAP K4, Photometrics).
Since the FOV is \( \sim 136 \, \mu m \) (limited by the \( \sim 15\)-mm image sensor, and the \( \sim 111 \times \) magnification), while the total length over which the wavelengths are dispersed in the focal plane exceeds 1 mm, the objective is scanned along the dispersion direction to cover the whole focal plane at each wavelength (11 images captured for each wavelength). These images were then combined to form the full intensity distribution at each wavelength. The measurements were performed at 11 wavelengths (760 nm to 860 nm, 10-nm steps) to form the results shown in Figs. 2.4b, 2.4c, Fig. 2.A6, and Fig. 2.A7. The measurements were also performed at a second set of wavelengths (761.25 nm, 811.25 nm, and 861.25 nm). These results are summarized in Fig. 2.4d and Fig. 2.4e for TE and TM polarizations. The resolution [Fig. 2.A8] was estimated by finding the FWHM at each wavelength in addition to the displacement rate of the focus center along the \( y \) direction by changing the wavelength.

The setup was slightly changed for measuring the focusing efficiencies. The input beam was partially focused by a lens (\( f = 10 \) cm) such that all the beam power passed through the input aperture (with a \( \sim 400 \, \mu m \) FWHM). In addition, the camera was replaced by a photodetector and a pinhole with a diameter of 3.5 mm in front of it to measure the focused power. The pinhole, corresponding to a \( \sim 31-\mu m \) area in the focal plane, allows only for the in-focus light to contribute to the efficiency. The efficiency is then calculated at each wavelength by dividing these measured powers by the total power tightly focused by a 10-cm focal length lens that was imaged onto the power meter using the same microscope (i.e., by removing the spectrometer and the pinhole).

The experimental setup for capturing the sample spectra is almost identical to Fig. 2.A5b, with the only difference of the polarizer being replaced by the sample of interest, and an 840-nm short-pass filter inserted before the sample. The light source was also replaced by a supercontinuum laser (Fianium Whitelase Micro, NKT Photonics).

**Angular response measurement**

To measure the angular response of the device we used the setup shown in Fig. 2.A9c, equipped with a rotating stage with 0.1° precision in the \( x-z \) plane and 0.002° in the \( y-z \) plane. The collimator (connected to the fiber coming from the source) was mounted on this rotating stage, where the folded spectrometer was exactly located at its center. The incident angles were adjusted accordingly for 0°, ±0.3°, ±0.6°, ±1° angles. As can be observed in Fig. 2.A9a, the focal spots did not change much in size as the angle is varied between -1° to +1° in \( x \)-direction. For measuring the tilt
angle in the y-z plane, the distance from the collimator to the device was measured to be 280 mm. In order to impose ±0.15° tilt in the y direction, the mounted collimator level is raised or lowered by 0.73 mm, and its tilt was adjusted accordingly such that the beam hits the center of the input aperture. As shown in Fig. 2.A9b, such a tilt in input incident angle does not degrade the spectral resolution of 1.25 nm.

**Design of the spectrometer with high throughput**

To further demonstrate the capabilities of the platform, we have designed a second spectrometer with significantly increased throughput. In order to achieve higher throughput, a larger input aperture is required, so the slab thickness was increased to 2 mm to give more freedom on the non-overlapping condition for the metasurfaces. The design, as shown in Fig. 2.A10, has a 2.5 mm input aperture. To further improve the throughput, the acceptance angle of the device was increased. To achieve this goal, we took an approach similar to the fabricated spectrometer with the difference of adding extra phase terms to the input diffraction grating. This helps with orienting the focuses on the image plane for different incident angles, as well as relaxing the condition for focusing in the x-direction. This in turn allows for increasing the input incident angle to ±15° degrees. The phase profile coefficients for metasurfaces 1 to 3 in Fig. 2.A10 are given in Table 2.A3. In the final design, the power is distributed in an area close to 200 μm wide in the x-direction in the focal plane, instead of a diffraction-limited focus. According to the intensity profiles shown in Fig. 2.A10b, the device can distinguish between wavelengths spaced by 0.5 nm both at the center wavelength of 810 nm, and also at the side wavelengths of 760 nm and 860 nm. Based on the angular response of the device in the x-z and y-z planes, and also the input aperture size of the device, an etendue of around ~13000 Srμm² is estimated.
2.3 Appendix: Additional figures and tables for the metasurface spectrometer

Figure 2.A1: Single-post periodic lattice simulations. (a) Schematic of a rectangular post on top of a fused silica substrate, showing the post dimensions. The nano-posts are capped by a 2-μm-thick layer of SU-8, and backed by a reflective gold layer. (b) Simulated reflection phase under TE illumination with 33.9° incident angle. (c) Derivative of the phase with respect to the wavelength calculated and plotted versus the height ($h$) and width of the post ($D_x = D_y = D$). The nano-post height that provides full $2\pi$ phase coverage with high reflectance while minimizing variation of the phase derivative is found to be $h=395$ nm (the red line). (d) Reflectance as a function of post-width and height.
Figure 2.A2: Reflection phase variation versus wavelength. (a) Reflection phase for TE-polarized light from a uniform array of meta-atoms corresponding to the black curves in Fig. 2.3 versus wavelength. As depicted by the vertical axis on the right, the simulation for each wavelength is performed under an incident angle corresponding to the deflection angle of the input blazed grating at that wavelength. (b) Same as a for TM polarization.

Figure 2.A3: Grating design curves and deflection efficiencies. (a) Post width versus reflection phase for 395-nm-tall posts on a square lattice with a 250-nm lattice constant. The red stars correspond to the nano-post sizes found from this graph that have the highest deflection efficiency over the bandwidth. (b) TE and TM polarization deflection efficiency curves versus wavelengths for the initial (i.e., the red stars in a) and optimized nano-post dimensions given in Table 2.A2.
Figure 2.A4: The key fabrication steps. A 395-nm-thick layer of α-Si is deposited on a 1-mm-thick fused silica substrate using PECVD. The metasurface pattern is generated with electron beam lithography, negated and transferred to the α-Si layer via lift-off and dry-etching processes. Both sides are covered with and SU-8 layer, and the input and output apertures are defined through photolithography and lift-off.

Figure 2.A5: Measurement setups. (a) Schematics of the measurement setup used for device characterization. (b) Schematics of the setup used to measure the focusing efficiencies. (c) An optical image of a part of the actual measurement setup showing the device, the objective lens, and the tube lens.
**Figure 2.A6: Focal plane intensity profiles.** (a) Two-dimensional intensity profiles measured at several wavelengths ($y_c$ is the center position of each profile) under illumination with TE-polarized light, and (b) TM-polarized light. (c) The position of the center of the focal spot along the dispersion direction, $y$, versus wavelength. The symbols represent the measured data, and the solid line is an eye guide. The scale bars are 20 μm.
Figure 2.A7: Intensity distribution profiles on logarithmic scales. (a) Same information as Fig. 2.4b and 2.4c of the main text, plotted on logarithmic scale for TE polarization, and (b)) for TM polarization.

Figure 2.A8: Measured spectral resolution versus wavelength. Spectral resolution estimated using the measured focal spot FWHM and the displacement rate of the focal spot with changing the wavelength. The average resolution is 1.22 nm for both polarizations. The symbols show the measured data and the solid lines are eye guides.
Figure 2.A9: Measured angular response of the device for polar angle variation with respect to 0 angle in x-z and y-z planes. (a) Angular response of the device for different tilted incident angles between -1° to +1° in the x-z plane. (b) Angular response of the device for ±0.15° tilted incident angles in the y-z plane. (c) Experimental setup used for measuring the angular response. The scale bars are 10 μm.
Figure 2.A10: Ray-optics design and simulation results of an extended-throughput folded spectrometer. (a) Ray tracing simulation results of the extended-throughput folded spectrometer, shown at three wavelengths in the center and two ends of the band. The system consists of three metasurfaces optimized to separate different wavelengths of the light and focus them on the focal plane. (b) Simulated intensity distribution for two wavelengths separated by 0.5 nm around three different center wavelengths of 760 nm, 810 nm, and 860 nm at 4 different incident angles of 0°, 5°, 10° and 15°. The intensity distributions show that wavelengths separated by 0.5 nm are theoretically resolvable for all aforementioned incident angles. The scale bars are 15 μm.
Table 2.A1: Phase profile coefficients in terms of \( [\text{rad/mm}^{m+n}] \) for metasurfaces I and II

<table>
<thead>
<tr>
<th>Metasurface</th>
<th>I (R=525.0 ( \mu )m)</th>
<th>II (R=582.5 ( \mu )m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( a_{x^2y^0} )</td>
<td>-4.02</td>
<td>-3.91</td>
</tr>
<tr>
<td>( a_{x^2y^1} )</td>
<td>-2.08</td>
<td>-3.70</td>
</tr>
<tr>
<td>( a_{x^2y^1} )</td>
<td>0.47</td>
<td>-0.68</td>
</tr>
<tr>
<td>( a_{x^2y^3} )</td>
<td>0.20</td>
<td>-0.24</td>
</tr>
<tr>
<td>( a_{x^4y^0} )</td>
<td>(-5.68 \times 10^{-4})</td>
<td>(6.26 \times 10^{-3})</td>
</tr>
<tr>
<td>( a_{x^2y^2} )</td>
<td>(7.55 \times 10^{-3})</td>
<td>0.021</td>
</tr>
<tr>
<td>( a_{y^4y^0} )</td>
<td>(2.36 \times 10^{-3})</td>
<td>(6.34 \times 10^{-3})</td>
</tr>
<tr>
<td>( a_{y^6} )</td>
<td>(1.93 \times 10^{-4})</td>
<td>(-2.48 \times 10^{-4})</td>
</tr>
<tr>
<td>( a_{y^8} )</td>
<td>(-3.22 \times 10^{-6})</td>
<td>(4.82 \times 10^{-6})</td>
</tr>
<tr>
<td>( a_{y^{10}} )</td>
<td>(-5.886 \times 10^{-9})</td>
<td>(-2.974 \times 10^{-9})</td>
</tr>
</tbody>
</table>

Table 2.A2: Optimized grating post sizes [nm] \((D_x, D_y)\)

<table>
<thead>
<tr>
<th>Optimization</th>
<th>Diffraction efficiency</th>
<th>Particle swarm optimization</th>
</tr>
</thead>
<tbody>
<tr>
<td>( D_{x1} )</td>
<td>93.4</td>
<td>68</td>
</tr>
<tr>
<td>( D_{y1} )</td>
<td>93.4</td>
<td>134</td>
</tr>
<tr>
<td>( D_{x2} )</td>
<td>117</td>
<td>115.2</td>
</tr>
<tr>
<td>( D_{y2} )</td>
<td>117</td>
<td>119.6</td>
</tr>
<tr>
<td>( D_{x3} )</td>
<td>132.8</td>
<td>147.4</td>
</tr>
<tr>
<td>( D_{y3} )</td>
<td>132.8</td>
<td>151.2</td>
</tr>
<tr>
<td>( D_{x4} )</td>
<td>155.4</td>
<td>137.8</td>
</tr>
<tr>
<td>( W_{y4} )</td>
<td>155.4</td>
<td>178.8</td>
</tr>
</tbody>
</table>
Table 2.A3: Phase profile coefficients of the increased throughput design (Fig. 2.A10) in terms of \([\text{rad/mm}^{m+n}]\)

<table>
<thead>
<tr>
<th>Metasurface</th>
<th>I (R=2.5 mm)</th>
<th>II (R=1.5 mm)</th>
<th>III (R=2.15 mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a_{x^0y^0})</td>
<td>-0.44</td>
<td>-3.13</td>
<td>-2.45</td>
</tr>
<tr>
<td>(a_{x^2y^2})</td>
<td>-2.98</td>
<td>1.91</td>
<td>-2.69</td>
</tr>
<tr>
<td>(a_{x^2y^4})</td>
<td>0.015</td>
<td>0.027</td>
<td>-0.067</td>
</tr>
<tr>
<td>(a_{x^2y^3})</td>
<td>0.035</td>
<td>-0.04</td>
<td>-0.06</td>
</tr>
<tr>
<td>(a_{x^4y^0})</td>
<td>(-1.5 \times 10^{-5})</td>
<td>(-8.2 \times 10^{-4})</td>
<td>(3.56 \times 10^{-4})</td>
</tr>
<tr>
<td>(a_{x^2y^5})</td>
<td>(-6.16 \times 10^{-5})</td>
<td>(-0.72 \times 10^{-4})</td>
<td>(-1.73 \times 10^{-5})</td>
</tr>
<tr>
<td>(a_{x^2y^7})</td>
<td>(-3.1 \times 10^{-4})</td>
<td>(3.8 \times 10^{-5})</td>
<td>(-9.14 \times 10^{-4})</td>
</tr>
<tr>
<td>(a_{x^6})</td>
<td>(1.42 \times 10^{-7})</td>
<td>(1.12 \times 10^{-5})</td>
<td>(6.71 \times 10^{-7})</td>
</tr>
<tr>
<td>(a_{x^8})</td>
<td>(-2.26 \times 10^{-10})</td>
<td>(-5.26 \times 10^{-8})</td>
<td>(-1.72 \times 10^{-9})</td>
</tr>
<tr>
<td>(a_{x^{10}})</td>
<td>(-5.886 \times 10^{-13})</td>
<td>(7.58 \times 10^{-11})</td>
<td>(1.62 \times 10^{-12})</td>
</tr>
</tbody>
</table>
MINIATURIZED HYPERSPECTRAL IMAGERS

The material in this chapter was in part presented in [134, 172].

3.1 Hyperspectral imager with folded metasurface optics

Hyperspectral imaging is a key characterization technique used in various areas of science and technology. Almost all implementations of hyperspectral imagers rely on bulky optics including spectral filters and moving or tunable elements. Here, we propose and demonstrate a line-scanning folded metasurface hyperspectral imager (HSI) that is fabricated in a single lithographic step on a 1-mm-thick glass substrate. The HSI is composed of four metasurfaces, three reflective and one transmissive, that are designed to collectively disperse and focus light of different wavelengths and incident angles on a focal plane parallel to the glass substrate. With a total volume of $8.5 \text{ mm}^3$, the HSI has spectral and angular resolutions of $\sim 1.5 \text{ nm}$ and $0.075^\circ$, over the 750 nm–850 nm and $-15^\circ$ to $+15^\circ$ degree ranges, respectively. Being compact, low-weight, and easy to fabricate and integrate with image sensors and electronics, the metasurface HSI opens up new opportunities for utilizing hyperspectral imaging where strict volume and weight constraints exist. In addition, the demonstrated HSI exemplifies the utilization of metasurfaces as high-performance diffractive optical elements for implementation of advanced optical systems.

Introduction

Hyperspectral imaging, originally developed and utilized in remote sensing [173, 174], is a very powerful technique with applications in numerous areas of science and engineering such as archaeology [175], chemistry, medical imaging [176], biotechnology, biology [177], bio-medicine, and production quality control [178, 179]. In general, a hyperspectral imager (HSI) captures the spectral data for every point in an image. Therefore, the hyperspectral data for a 2D image is a 3D cube in which the first two dimensions are the spatial directions and the third one represents the spectrum (i.e., the imager records the function $I(x, y, \lambda)$).

Several methods and HSI platforms have been developed to acquire the 3D data cube using the existing 2D image sensors. One category of HSIs use tunable band-pass filters that can sweep through the desired spectral band [180, 181]. In
these devices, a 2D image is captured at each step in the scan, recording the optical power within the filter bandwidth. The required spectral scanning setups usually rely on a fine tuning mechanism that might not be fast or compact enough for many applications. A significant effort has been made to develop HSIs with faster and more compact spectral scanning schemes [182, 183] and lower aberrations [184]. While acousto-optical and liquid crystal tunable filters provide solutions for fast spectral scanning, their low-throughput (under 50%) is still a disadvantage of these tunable filters [181].

Another class of devices, snapshot HSIs, acquire the 3D data cube in a single shot without the need for a scanning mechanism [185, 186]. However, they generally require heavy post-processing and rely on some level of sparsity in the spectral and/or spatial content of the image [187] as they are, in essence, compressive sensing methods. While their higher data rates and speeds make them suitable for recording transient scenes [188], they generally suffer from low signal to noise ratios (SNR), and require significant computational resources [187]. Snapshot image mapping spectrometers (IMS), based on the idea of image slicing and dispersing each slice to obtain the spectral information and reconstruct the 3D data cube, work well only for low spatial resolution applications [189]. Additionally, the image mapper which is the primal part of IMS hyperspectral imaging systems needs to cut the scene with a high precision and often are not compact [190].

A different group of HSIs are based on spatial scanning, and require a relative displacement of the HSI and the object of interest (i.e., either the object or the HSI is moved in space) [191]. The spatial scanning is either done pixel by pixel (point scanning/whisk-broom) [192] or line by line (push-broom) using a slit in front of the HSI [193]. The whisk-broom technique requires 2D spatial scanning which results in longer acquisition times in comparison to the push-broom method. Thus, its applications are mostly limited to cases like confocal microscopy where measuring one point at a time while rejecting the signal from other points is of interest [178]. The push-broom HSIs [194] are faster and better-suited for applications such as air- and spaced-based hyperspectral scanning where the whole scene of interest might not be at hand at once [195]. One advantage of push-broom HSIs is that a large number of spectral-bands are captured without the burdensome post processing that is generally required for snapshot HSIs. Moreover, push-broom HSIs generally provide higher SNRs and better angular resolution compared to the snapshot ones [193]. Other approaches that indirectly obtain the 3D data cube, such as interferometric Fourier
transform spectroscopic imaging [196], in general rely on bulky and complicated optical setups, and are not well suited for compact and low-weight systems [197].

A common challenge with almost all of the mentioned platforms is their compact, robust, and low-weight implementation, limited by the requirement for relatively complicated optical systems and reliance on mostly bulky conventional optical elements. In recent years, dielectric optical metasurfaces have overcome some of the limitations faced by the conventional optical elements [1, 47, 198, 199]. Their ability to control the phase [52, 200–202], phase and polarization [25], and phase and amplitude [203, 204] of light on a sub-wavelength scale and in compact form factors has made them very attractive for the implementation of compact optical systems [130, 131, 205]. In addition, the additional available degrees of freedom in their design allow for devices with enhanced control [26, 206, 207] that are otherwise not feasible.

Here, we propose and experimentally demonstrate a push-broom HSI based on the folded metasurface platform [133]. In this device, three reflective and one transmissive dielectric metasurfaces, which are monolithically fabricated on a glass substrate in a single lithographic step, disperse and focus light for various incident angles and wavelengths on a image plane parallel to the substrate. Working in the 750 nm–850 nm wavelength range, the metasurface HSI provides more than 70 resolved spectral points in a 8.5 mm$^3$ volume. Spatially, the HSI has an angular resolution of $\sim 0.075^\circ$ and distinguishes about 400 angular directions in the $\pm 15^\circ$ range. The compact form factor, low weight, and high level of integrability of the metasurface HSIs make them very attractive for utilization into devices like consumer electronics, and more generally applications where there are stringent volume and weight limitations in addition to high performance requirements.

**Concept and design**

A push-broom HSI is schematically shown in Fig. 3.1a. The HSI captures a 1D spatial image along the direction $\theta$, measuring the spectrum at each pint along the line independently. The full 3D data cube can be formed by scanning the object in front of the HSI. Figure 3.1b schematically shows a folded metasurface push-broom HSI. The HSI includes one transmissive and three reflective metasurfaces patterned on one side of a 1-mm-thick fused silica substrate with gold mirrors on both sides. Light enters the HSI through an input aperture in one of the gold mirrors and is deflected into the substrate and vertically dispersed by the first metasurface (which acts as first-order blazed grating). The other two reflective metasurfaces together with
Figure 3.1: Concept of a push-broom folded metasurface HSI. (a) Schematic illustration of a push-broom hyperspectral imager. The spectral content is captured sequentially for different cross-sections of the object. (b) The proposed scheme for a folded metasurface hyperspectral imager. The device includes multiple reflective and transmissive metasurfaces each performing a different optical function. Light enters the device through an input aperture, interacts with the reflective metasurfaces while it is confined inside the substrate by the two gold mirrors, and exits the output aperture that has a transmissive metasurface built into it. Different wavelengths are dispersed in the vertical direction ($\lambda$), and various input angles are focused to different horizontal points.

The transmissive one focus light with different wavelengths and horizontal incident angles to diffraction-limited spots on the detector array plane that is parallel to the substrate. In this configuration, the transmissive metasurface, which is defined in the same lithography step as the reflective ones, simultaneously acts as the output aperture.

Figure 3.2a illustrates ray tracing simulations of the designed HSI. The first metasurface acts as a blazed-grating with a 1-μm period, dispersing the collimated light coming from different angles, into angles centered at 33.46 degree (in the y-z plane) at the center wavelength of 800 nm. The phase profiles of the other three metasurfaces are optimized to provide near diffraction-limited focusing for the 750 nm–850 nm spectral and ±15 degree spatial range on a focal plane parallel to the substrate. For ease of measurements, the focal plane is designed to be 1 mm outside the substrate and parallel to it. The detailed design process of the three metasurfaces along with their phase profile coefficients are presented in Appendix 3.3.

The three reflective metasurfaces are implemented using a platform similar to the one used in [133]. The metasurfaces consist of $\alpha$-Si nanoposts with rectangular
Figure 3.2: Ray-optics design and images of the fabricated devices. (a) Ray tracing simulation results of the folded metasurface HSI, shown for three wavelengths and a 0-degree incident angle. The system consists of 4 metasurfaces, first of which acts as a blazed grating dispersing different wavelengths. The following three metasurfaces (I, II, and III) are optimized to correct aberrations and focus the rays on the image plane for the desired wavelengths (750 nm–850 nm) and angles (-15 to +15 degrees). (b) An optical microscope image showing six HSIs on the same chip (left) and zoomed-in images of the metasurfaces that comprise one HSI (right). The images were captured before depositing the second gold layer. Scale bars: 1 mm (left) and 500 μm (right). (c) Scanning electron micrographs of parts of the fabricated metasurfaces. Scale bars: 1 μm.

cross-sections, resting on a 1-mm-thick fused silica substrate and capped by a 2-μm-thick SU-8 layer. The nanoposts’ height and lattice constant are 395 nm and 246 nm, respectively, and a gold layer is deposited on the SU-8 layer to make the metasurfaces reflective. For the transmissive metasurface (III), the height and lattice constant are 600 nm and 250 nm, respectively, and there is no gold mirror. The nanopost heights are in both cases chosen to achieve full $2\pi$ phase coverage, while minimizing the variation in the wavelength derivative of phase to keep the diffraction efficiency high over the whole bandwidth [206]. In addition, the lateral dimensions of the nanoposts are selected to make the metasurfaces polarization independent for an operation angle of 33.46 degrees [133]. We also confirmed that the metasurfaces are almost polarization independent for the entire operation bandwidth and angular range (see Figs. 3.A2 and 3.A3 for simulated transmission and reflection phases). The blazed grating has a 1-μm period consisting of 4 nano-posts located on a lattice with a 250-nm lattice constant. The optimization procedure used for its design is similar to the approach taken in [133]. An initial structure was found using the high-NA high-efficiency design method [124], and was then further optimized using the particle swarm optimization technique.
Experimental results

The metasurface HSI fabrication process was started with depositing a 600-nm-thick layer of $\alpha$-Si on the fused silica substrate. Since the reflective and transmissive metasurfaces have different thicknesses (i.e., 395 nm and 600 nm, respectively), areas of the sample containing the reflective metasurfaces were first etched down to 395 nm. The patterns for both reflective and transmissive metasurfaces were then defined in a single electron beam lithography step, eliminating the need for additional alignment procedures. In addition, to avoid over-etching the thinner parts (corresponding to the reflective metasurfaces), dry etching was performed in two separate steps and the thinner parts were protected by a photoresist during the second step. After the etch and removal of the mask, the metasurfaces were covered by a 2-$\mu$m-thick SU-8 layer. The input and output apertures were then defined in ~100-nm-thick gold layers deposited on both sides of the substrate. The mirrors on both sides were then covered by another 2-$\mu$m-thick SU-8 layer for protection. Optical microscope images of the devices before deposition of the second (left) and first (right) gold mirrors are shown in Fig. 3.2b. Scanning electron micrographs of parts of the fabricated metasurfaces before capping with the SU-8 layer are shown in Fig. 3.2c. See Appendix 3.3 for fabrication details.

To characterize the fabricated HSI, the input aperture of the device was illuminated by a collimated beam from a tunable continuous-wave laser, and its focal plane was imaged using a custom-made microscope [Fig. 3.3a]. The setup was designed to allow for the input beam to be rotated in the horizontal plane around the input aperture (see Fig. 3.A4a and Appendix 3.3 for details of the measurement setup). The input wavelength was tuned in the 750 nm to 850 nm range in steps of 10-nm, and the input angle was changed in 5° separations. The intensity distribution in the focal plane was captured using the microscope. Figure 3.3b shows the captured intensities at 750, 800, and 850 nm for multiple incident angles. For comparison, the simulated expected focal spots positions are given in Fig. 3.3c, highlighting the spots where the zoomed-in distributions are plotted in Fig. 3.3b. Measured intensity distributions at other wavelengths and angles are given in Fig. 3.A5.

Figure 3.3d shows the simulated full width at half maximum (FWHM) along the y direction, the simulated spectral resolution, and the measured FWHMs at several points. Figure 3.3e shows similar results for the angular resolution of the device and the FWHM along the x direction. The good agreement observed between the simulated and measured FWHMs denotes the nearly diffraction-limited operation of
the fabricated device. In addition, it can be seen from these results that the device has spectral and angular resolutions better than 1.4 nm and about 0.075 degree, respectively, across the whole bandwidth and for various angles. As a result, the demonstrated HSI can resolve more than 70 spectral and 400 angular points. In addition, we measured the focal spots for multiple sets of wavelengths in the range with 1.5-nm separations, and several sets of angles with 0.1-degree distances. The results (see Fig. 3.4a) confirm the resolving power of the HSI.

The HSI focusing efficiency was measured using an approach similar to [133], and its average value was found to be \( \sim 10\% \) (see Appendix 3.3 and Figs. 3.A4b and 3.A6 for the measurement details, setup, and measured efficiencies, respectively). The lower efficiency achieved here (in comparison to the spectrometer in [133]) is mostly because of using 4 metasurfaces instead of 3, especially as the transmissive metasurface has a large deflection angle of \( \sim 33 \) degrees in glass. The efficiency can be increased by using mirrors with higher reflectivities (like distributed Bragg reflectors), using anti-reflection coatings on the input and output apertures, and further optimizing the design and fabrication of the metasurfaces, especially the dispersive grating and the transmissive metasurface. The focusing efficiency reduces by \( \sim 36\% \) at the incident angles of ±15 degrees. The reduction in efficiency at higher incident angles compared to the normal one can be attributed to different factors. Primarily, the effective aperture seen by oblique beams is smaller at higher incident angles. Furthermore, the diffraction grating, which is optimized at normal incidence, does not operate as efficiently at other incident angles. Finally, the phase profiles of the metasurfaces, slightly deviate from their optimal profiles at oblique incident angles.

To demonstrate the capability of the folded metasurface HSI in recovering 3D hyperspectral data cubes we used it to image an object with spatially varying spectral information. To this end, we designed and fabricated an imaging target with narrow transmission dips over the 750 nm–850 nm wavelength range (see Appendix 3.3 and Fig. 3.A8 for more information on the design and fabrication of the target). As shown in Fig. 3.4a, a broadband source (covering 750 nm–850 nm) was used to illuminate the object. The imaging optics in Fig. 3.4a mapped different points along a horizontal line on the target to different incident angles on the HSI input aperture within its acceptance range (see Fig. 3.A4c for details of the measurement setup). To measure the hyperspectral content, the target was vertically moved and the focal plane of the HSI was imaged at each step of the movement. This results in a 3D
Figure 3.3: Simulated and experimental characterization results of the folded metasurface HSI. (a) Simplified schematic of the measurement setup. A collimated beam from a tunable laser illuminates the input aperture at various incidence angles, and the intensity distributions at the focal plane are captured. (b) Measured focuses on the image plane, shown for 3 wavelengths (750, 800, and 850 nm) and angles ranging from -15° to +15° with 5° increments. The insets show the zoomed-in version of intensity profiles (scale bars: 10 μm). (c) Simulated locations of the spots on the focal plane for wavelengths increasing from 750 nm (blue) to 850 nm (green) at 10-nm steps, and angles from -15° (left) to +15° (right) at 5° separations. The highlighted focal spots are magnified in the insets of b. (d) Simulated spectral resolution and vertical FWHM versus wavelengths, calculated at multiple incident angles. The measured FWHM values for three wavelengths are overlaid on the graph. (e) Simulated angular resolution and FWHM versus wavelengths calculated for multiple incident angles. Measured horizontal FWHM values at three wavelengths are overlaid on the graph.
data cube, with a 2D image corresponding to each horizontal line on the target (see Fig. 3.A7 for details). Using the data in Fig. 3.3c, the intensity maps can be converted to angular/spectral data for each line. The inset in Fig. 3.4a shows the expected transmission dip wavelengths over the target. Figure 3.4b shows the captured images of the target at four different wavelengths in the range. For comparison, we also imaged the target using a tunable laser (TL) by scanning the illumination wavelength through the range (see Fig. 3.A4d for the measurement setup and other details). The data obtained with the TL setup is shown next to the HSI images in Fig. 3.4b, denoting a good agreement between the results. In addition, Fig. 3.4c shows the captured spectra along a vertical and a horizontal cut in the target, using both the HSI and the TL setup. We mostly attribute the minor differences between the two sets of measurements to the speckle noise observed in the images obtained using the tunable laser. The speckle noise was caused by the residual degrees of spatial and temporal coherence of the tunable laser which persisted despite the use of a rotating diffuser.

**Conclusion**

We demonstrated a push-broom metasurface hyperspectral imager with a volume of 8.5 mm$^3$ and weighing less than 20 mg, resolving more than 70 spectral and 400 angular points in the 750 nm–850 nm and ±15 degrees range, respectively. The significant reduction in size and weight of the device, achieved through the folded architecture design, makes the device a promising candidate for applications where compactness, low-weight, and robustness are of primary importance. In addition, the possibility of mass producing many similar or different devices in a single lithographic step, make them easy and potentially low-cost to fabricate. More broadly, as one of the first demonstrations of a meta-system consisting of multiple reflective and transmissive metasurfaces performing a complicated optical function in a small form factor, this work paves the way towards better realization of potentials of metasurfaces in implementing advanced functional optical and optoelectronic systems for operation under strict constraints.
Figure 3.4: Hyperspectral imaging of a target with the folded metasurface HSI (M-HSI) and a tunable laser. (a) Simplified schematic of hyperspectral imaging a target with the HSI. The imaging optics maps different points along a horizontal line on the target to collimated beams with different incident angles on the HSI aperture. The image containing the spectral and angular data for the line is formed by the HSI in its focal plane and captured by a custom-built microscope. The inset shows the fabricated target, which has transmission dips with increasing center wavelength from bottom to the top. (b) Measured transmitted intensity of the object at four wavelengths captured by the metasurface HSI, compared with the ones measured using a tunable laser (TL). (c) Transmitted intensity profile of the object along the two cuts A and B, compared with the same results obtained using the tunable laser.
3.2 Random phase mask metasurface hyperspectral imager

In recent years, there has been a high interest in “snapshot” hyperspectral and/or multi-spectral imaging, where the 3D data-cube of the object or its spatial content at multiple spectral bands can be acquired in a single shot, but after further computational post-processing [178]. High data rates of capturing is thus enabled by snapshot hyper/multi-spectral imaging techniques making them highly recommended for applications which involve recording the transient scenes. Several approaches for snapshot hyper/multi-spectral imaging have been suggested and reported so far; among them Coded Aperture Snapshot Spectral Imaging (CASSI), Computed Tomography Imaging Spectroscopy (CTIS), Image Mapping Spectrometry (IMS) are well developed [185].

One drawback in the majority of snapshot imaging techniques is the requirement for massive computational processing to obtain the 3D data cube. In specific CTIS suffers from the missing cone problem [208, 209]. In CASSI, the coded aperture mask which is a random binary matrix limits the photon transmission to about 50% [186]. Furthermore, the 3D data cube recovery is done by a reconstruction algorithm which requires a compressible basis, that results in spectral limitations in resolving point sources near the scale of the optical resolution as well as being prone to the artifact imposed on the system measurement matrix [187]. IMS, which is based on the idea of image slicing and dispersing each slice to obtain the spectral information and reconstruct the 3D data cube, works well only for low spatial resolution applications [189]. The image mapper, which is the primal part of IMS hyperspectral imaging systems, needs to cut the scene in very high precision and are not usually compact [190]. Other approaches such as interferometric Fourier transform spectroscopic imaging [196], investigated in the literature, involves utilizations of bulky optical elements, which might be of some potential for field utilizations, but hinders their usage for applications requiring compact, fast, cost-effective, high-resolution imaging.

Dielectric metasurfaces on the other hand overcome the challenges faced by conventional optical systems in a compact form factor. In addition to controlling the phase and polarization of wave-fronts independent of each other, dielectric metasurfaces have facilitated the implementation of conventional and novel functionalities with high efficiencies, otherwise not feasible by traditionally-used bulky optical elements. Many of the complex, heavy and bulky optical systems such as those mentioned above, can be realized by means of dielectric metasurfaces, due to their unique ability
for monolithic integration with the well-known nano-fabrication techniques. Here, a computational hyper/multi spectral imaging scheme using a random metasurface phase mask is illustrated.

**Introduction**

Most of the optical imaging systems, unless corrected for, suffer from spectral aberrations. As an example, a perfect lens with zero geometrical aberration focuses various spectral bands at different distance from it. Although the image on the image plane is perfectly focused at the center wavelength, it blurs as the wavelength deviates from that wavelength. Consequently, spectral aberration is the cause for conventional non-achromatic imaging systems not be capable of imaging objects at multiple spectral bands. Thus, a class of multispectral imaging systems are introduced to operate based on scattering of light passing through a random medium which produces speckle patterns on the image plane [210]. No coded aperture (amplitude) mask is used in these schemes. In this case, the speckle patterns projected on the image plane are desired to be well de-correlated across spectral bands, so the spatial content of the object at each spectral band can be reconstructed by a simple decorrelation function, which can be either done in real space domain or the Fourier domain. The technique works well for recovery of the objects at small number of bands as in [210], but as the number of wavelength increases to tens or hundreds the approach becomes less effective. Hence, the successful reconstruction of multi-spectral objects from the two-dimensional speckle patterns requires further assumptions about the object such as sparsity, spatial spectral separations and also about the measurement matrix such as being high rank and/or low coherence [211]. Assuming some level of sparsity for the multispectral object of interest either in spatial or spectral forms, several computational schemes for reconstructing the object have been suggested in the literature [212–217].

Here we investigate how a random dielectric metasurface mask will generate speckle patterns on the image plane at different wavelengths in near infrared. In addition, the design of a random metasurface mask to create maximum decorrelation across the spectral band is studied. Using the generated speckle patterns several computational optics techniques to reconstruct the object along with their reconstruction accuracy will be discussed.

The setup for hyperspectral imaging of the object is depicted in Fig. 3.5a, where the object is placed at a distance from the random metasurface mask and the speckle patterns are projected on the image plane, subsequently captured by the lens and the
Figure 3.5: Hyperspectral imaging using a random metasurface phase mask. (a) Schematic of the setup consisting the metasurface random mask and projected speckles on the image plane. (b) Spectral reconstruction of the object from the captured speckle pattern.

camera. Using the reconstruction algorithm, explained later in details, the spatial content of the object at each wavelength can be recovered as shown in Fig. 3.5b. In order to achieve maximum spectral resolution, it is ideal that the speckle patterns of a point source placed at the object plane are de-correlated across all wavelengths.

Figure 3.6a shows a phase mask with uniform phase distribution in space (x, y). The point source object is placed at the distance of 1 mm from the mask and the mask is 1 mm far from the image plane. The intensity and phase of the field at the image plane are captured at 850 nm, shown in Fig. 3.6b. Based on the speckle patterns generated for all wavelengths on the image plane, the correlation coefficients are calculated and plotted as shown in Fig. 3.6c. As depicted in Fig. 3.6c there is a small correlation background across the spectrum, which is the indication of remaining residual correlations between the wavelengths in the range 700 nm-1 μm. The average decorrelation resolution in this case (where the decorrelation function reaches 50%) is estimated to be around 2-3 nm. As shown in Fig. 3.6b the field (intensity and phase) is not uniformly distributed on the image plane, in other words the degrees
Figure 3.6: **Optical properties of a uniform distribution phase mask.** (a) Phase distribution of the phase mask, uniformly distributed between 0-2π phases. (b) Correlation coefficients calculated for the speckle patterns at different wavelengths in the range of 0.7 to 1 μm. (c) The intensity and phase of the speckle pattern for a point source which is placed 1 mm away from the phase mask at λ =850 nm.

To achieve a phase mask distribution with maximum decorrelation across the range of wavelength, the algorithm similar to Gerchberg–Saxton (GS) algorithm is used, which starts from the uniform distribution in space. At each iteration the amplitude unity mask for the Fourier components having a maximum numerical aperture of 0.5 is applied. After a couple of iterations the phase mask converges to the one shown in Fig. 3.7a. The intensity and phase of the speckle pattern which is the point spread function (PSF) calculated at 850 nm are also demonstrated in Fig. 3.7b. As expected, the background correlation due to the remaining correlation across the range is now removed and the decorrelation resolution is calculated to be ~1 nm. In this case, the maximum number of resolved spectral bands are ~300 wavelengths.
Figure 3.7: Optical properties of a phase mask designed using the GS algorithm. (a) Obtained phase distribution of the mask after applying the GS algorithm. (b) Correlation coefficients calculated for PSFs at different wavelengths (c) The intensity and phase of PSF obtained from the designed phase mask at $\lambda = 850 \text{ nm}$.

The spatial resolution can also be calculated in the same manner for transversal $(x, y)$ and axial $(z)$ directions by calculating the cross-correlation function for point spread functions of the neighboring pixels at any wavelength. The spatial correlation function is plotted at $\lambda = 850 \text{ nm}$ in Fig. 3.8a. The half width at half maximum (HWHM) which can be estimated as the spatial resolution is $\sim 500 \text{ nm}$. The correlation function for two depths of the object placed at $1 \text{ mm}$ and $0.99 \text{ mm}$ away from the phase mask are also depicted in Fig. 3.8b. There is a small correlation of about 10-20% between the PSFs at these two depths. Hence, the axial decorrelation resolution can is estimated to be less than $10 \mu \text{m}$.

By analyzing the spectral and spatial decorrelation resolution of the system, the hyperspectral datacube of any defined object can be reconstructed from the obtained speckle patterns. As an example, a hyperspectral object having the spectral content as shown in Fig. 3.9a is defined, where each digit (0 to 9) is located at one specific
Figure 3.8: Spatial and spectral correlation of designed phase mask. (a) Spatial correlation function calculated at $\lambda=850$ nm for neighboring pixels on the object plane. (b) Spectral correlation function calculated corresponding to PSFs of two point sources located at two different depths (First mask at the distance of 1 mm from the mask and the second one at 0.99 mm).

wavelength and the spectral spacing between two consecutive digits is 30 nm. The objects at transition wavelengths corresponding to two consequent digits are defined using the linear interpolation.

Figure 3.9: Spectral content of the object used for hyperspectral imaging. The spectral content of a sample object defined at wavelength across the range of 0.7-1 $\mu$m (left). At wavelengths between the two consecutive digits the linear interpolation of the digits are used. The scheme showing the object and image plane positions, $Z_1 = Z_2 = 1$ mm (right).

Reconstruction

By defining the spectral object and placing it at the distance of 1 mm from the designed phase mask as shown in Fig. 3.9b, the speckle pattern of the object on the image plane can be captured. There are several techniques for reconstructing the object from the obtained speckle patterns:
The first and the simplest approach for reconstruction of a multi-spectral object is deconvolving the captured speckle pattern by the PSFs obtained at each wavelength (see Appendix 3.4 and Fig. 3.A9 for further details). This technique is used in [210] for visible light imaging using 3 beams. The approach however works well only for small number of spectral bands and fails when the number of wavelengths exceeds some limits (see Appendix 3.4 section for further details).

The second technique for object reconstruction uses compressive sensing, as a form of optimization problem, and is based on the assumption that the object of interest has some degrees of sparsity. Even if the object of interest is not itself sparse, there is a sparse basis that the object can be represented in it by means of a representation matrix and a sparse vector. Furthermore, most of the real objects used in nature for multi-spectral imaging in optics have finite spectral signatures and are low rank. Thus, the assumption of having both of the properties, i.e. being sparse and low-rank, makes the compressive sensing problem simpler and faster to solve [216]. Various schemes of compressive sensing hyperspectral imaging formulations exist in the literature [212–217], among which we utilized the one based on joint sparse and low-rank recovery of data-cubes (see Appendix 3.4 for further details). In addition to the way an optimization problem is defined, the algorithm to solve the problem requires to be fast and efficient in speed and also complexity. Different packages of convex optimization solvers such as cvx, ℓ1-magic are available online; among them l1-magic was able to reconstruct a data-cube with the dimension of \((51 \times 51 \times 21)\) for \((x, y, \lambda)\) in less than half an hour on a PC computer. The object at multiple wavelengths and the reconstructed image are shown in Fig. 3.10. The speed can be further accelerated by using novel optimization algorithms.

Figure 3.10: Object reconstruction at different wavelengths. The original and reconstructed objects using ℓ1-magic optimization solver at different wavelengths shown above each column. The width of each image is 100 μm.
Conclusion

A compact hyperspectral imaging concept using a metasurface phase mask has been introduced. The scheme has a spectral resolution of 1 nm, resolving ~300 spectral bands in the near-IR range (0.7-1 μm). Two reconstruction techniques for obtaining the hyperspectral data-cube of objects were studied. The reconstruction algorithm works well for joint sparse and low rank objects. Further development of the optimization solver will result in faster reconstruction of the hyperspectral data-cubes.

3.3 Appendix: Additional information and discussion for folded metasurface hyperspectral imager

Simulation and design

The system level design and simulation of the folded metasurface hyperspectral imager is performed using Zemax OpticStudio, where metasurfaces were defined as phase-only diffractive elements. The diffraction grating was modeled as a phase-surface with a linear phase along the dispersion direction (i.e., y direction). The blazed grating, with a period of 1 μm, deflects the normal incident light into deflection angles of 31.14° and 35.88° at 750 nm and 850 nm, respectively. The phase profiles of metasurfaces I, II and III were defined as polynomial expansions in Cartesian coordinates, Σn,mam,nx ny (named Binary 1 surfaces in Zemax OpticStudio). The am,n expansion coefficients were optimized to reduce geometrical aberrations for input angles within the range of -15° to +15° and wavelengths in the range of 750 to 850 nm. This was done by minimizing the geometric radii of image-plane focuses corresponding to each angle and wavelength, which finally leads to maximizing the angular and spectral resolution of the designed hyperspectral imager. The optimized coefficients for metasurfaces I, II, and III, as shown in Fig. 2 of the main manuscript, are given in Table 3.A1.

The phased profiles of metasurfaces I, II, and III are plotted in Fig. 3.A1a. The point spread functions (PSFs) for three wavelengths (750 nm, 800 nm, and 850 nm) and angles of 0°, 5°, 10°, and 15° are calculated for two cases. In the first case, shown in Fig. 3.A1b, the PSFs for two beams with the same incident angle, but separated spectrally by 1.5 nm centered around the specified wavelengths are plotted, and seen to be well resolved. In the second case, plotted in Fig. 3.A1c, PSFs for two beams with the same wavelength but separated by 0.1° around their corresponding incident angle show two resolved peaks. These verify the upper limits for spectral
and angular resolutions of $\sim 1.5\,\text{nm}$ and $\sim 0.1^\circ$, respectively. Due to the symmetry of the metasurfaces along the $x$-axis (see Fig. 3.A1a), the same conclusion holds true for negative incident angles. Furthermore, although only three different spectral points and four angular points were shown in Figs. 3.A1b and c, it is worth mentioning that the design is close to diffraction-limited for all wavelengths and angles within the design range.

**Table 3.A1: Optimized phase profile coefficients in terms of $[\text{rad/mm}^{m+n}]$**

<table>
<thead>
<tr>
<th>Metasurface</th>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>$2.1 \times 1.1$</td>
<td>$2.48 \times 0.8$</td>
<td>$2.6 \times 0.96$</td>
</tr>
<tr>
<td>$a_{x2,0}$</td>
<td>-5.087</td>
<td>-2.898</td>
<td>-3.513</td>
</tr>
<tr>
<td>$a_{x0,2}$</td>
<td>-9.197</td>
<td>7.229</td>
<td>11.137</td>
</tr>
<tr>
<td>$a_{x2,1}$</td>
<td>0.052</td>
<td>0.029</td>
<td>-0.872</td>
</tr>
<tr>
<td>$a_{x4,0}$</td>
<td>0.234</td>
<td>-1.494</td>
<td>1.422</td>
</tr>
<tr>
<td>$a_{x0,3}$</td>
<td>$1.84 \times 10^{-3}$</td>
<td>$-4.207 \times 10^{-5}$</td>
<td>0.011</td>
</tr>
<tr>
<td>$a_{x2,2}$</td>
<td>$-2.064 \times 10^{-3}$</td>
<td>$6.308 \times 10^{-3}$</td>
<td>$-6.571 \times 10^{-3}$</td>
</tr>
<tr>
<td>$a_{x4,1}$</td>
<td>$-4.002 \times 10^{-3}$</td>
<td>0.140</td>
<td>-0.043</td>
</tr>
<tr>
<td>$a_{x0,4}$</td>
<td>$-1.153 \times 10^{-4}$</td>
<td>$4.355 \times 10^{-4}$</td>
<td>$-1.397 \times 10^{-3}$</td>
</tr>
<tr>
<td>$a_{x2,3}$</td>
<td>$2.754 \times 10^{-4}$</td>
<td>$-3.081 \times 10^{-3}$</td>
<td>$2.947 \times 10^{-3}$</td>
</tr>
<tr>
<td>$a_{x4,3}$</td>
<td>$6.19 \times 10^{-5}$</td>
<td>-0.020</td>
<td>$6.42 \times 10^{-5}$</td>
</tr>
</tbody>
</table>

**Figure 3.A1: Optimized phase profiles and simulated PSFs for the HSI.** (a) Phase profile of metasurfaces I, II, and III. (b) Simulated PSFs for two incident beams with the same incidence angle, but separated spectrally by 1.5 nm (scale bars: 10 $\mu$m). (c) Simulated PSFs for two beams with the same wavelength, but separated angularly by $0.1^\circ$ around their corresponding incident angles (scale bars: 10 $\mu$m).

The rigorous coupled wave analysis technique was used to simulate the nano-posts [218]. The unit cells for reflective (I and II) and transmissive (III) metasurfaces are shown in Fig. 3.A2a and Fig. 3.A2g, respectively. The design method utilized to
achieve high-efficiency and polarization independent metasurfaces with minimum variation of phase derivative with wavelength is similar to the approach taken in [133]. First, a uniform array of $\alpha$-Si nano-posts were illuminated under an angle of 33.46° at the center wavelength of 800 nm, and the reflection (for metasurfaces I and II) or transmission (for metasurface III) phases and amplitudes were extracted. The height of nano-posts were then chosen such that the variation of the derivative of phase with respect to wavelength for various side lengths of nano-posts were minimized, while achieving a full $2\pi$ phase coverage and high efficiency. Considering the results of Figs. 3.A2d to 3.A2f and 3.A2j to 3.A2l, the optimal height values were chosen to be 395 nm (for the reflective metasurfaces I and II) and 600 nm (for the transmissive metasurface III). The lattice constants were chosen to be 246 nm (for reflective metasurfaces) and 250 nm (for the transmissive metasurface) in order to satisfy the sub-wavelength condition and avoid higher order diffractions [50].

To realize polarization-independent metasurfaces, the reflection and/or transmission phases were extracted for the various dimensions ($D_x$ and $D_y$) at the center wavelength of 800 nm and an illumination angle of 33.46°. The design curves were then generated through determining the path in the $D_x$-$D_y$ plane where reflection/transmission phases are almost equal for TE and TM polarizations. These curves are shown in Figs. 3.A2b and 3.A2c for the reflective metasurfaces, and Figs. 3.A2h and 3.A2i for the transmissive metasurface.

The hyperspectral imager was designed to behave similarly for the desired range of angles ($\pm 15^\circ$) and wavelengths (750–850 nm). To this end, metasurfaces need to impose almost similar reflection/transmission phases for different incident angles and wavelengths. The reflection phases depicted in Fig. 3.A3 demonstrate this insensitivity to angle for the reflective metasurfaces. The smoothness of the phase across the incident angles ($0^\circ$–$15^\circ$) verifies the almost angle-independent behavior of the nano-posts within our designed angular range at the center wavelength of 800 nm. Using the mentioned design approach, we have previously demonstrated that nano-posts have a minimal wavelength dependence as well [133].

For designing the blazed grating, the same height of $\alpha$-Si thickness as the reflective metasurfaces (395 nm) was chosen. The lattice constant was selected to be 250 nm, such that each period contains four nano-posts. The dimensions of these four nano-posts, optimized (see [133]) to achieve maximum diffraction efficiency over the entire bandwidth, are given in Table 3.A2.
Figure 3.A2: Unit cells and design curves showing dimensions of nano-posts to realize polarization independence. (a) Unit cell used to implement the reflective metasurfaces. (b) Reflection phase of nano-posts for an incident angle of 33.46° versus their side lengths for TE and (c) for TM polarizations. The black solid curves show the path in the $D_x$-$D_y$ plane where nano-posts have the same reflection phase for TE and TM polarizations. (d) Reflection phase of square nano-posts with square cross-sections plotted versus their side length and height for TE polarization at the center wavelength of 800 nm. (e) Phase derivative of square cross-section nano-posts with respect to wavelength plotted versus their side length and height. The red line shows the height (395 nm) where the phase derivative is minimal. (f) Reflectance plotted versus side-lengths and heights of nano-posts shown in a. (g) Unit cell used for the transmissive metasurface. (h-i) Same as b-c plotted for transmission phase. (j-l) Same as d-f plotted for the transmission function. The black lines show the height (600 nm) where phase derivative variation is minimal.
Table 3.A2: Optimized post sizes \((D_x, D_y)\) for the blazed diffraction grating

<table>
<thead>
<tr>
<th>Posts</th>
<th>(D_{x1})</th>
<th>(D_{y1})</th>
<th>(D_{x2})</th>
<th>(D_{y2})</th>
<th>(D_{x3})</th>
<th>(D_{y3})</th>
<th>(D_{x4})</th>
<th>(D_{y4})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Size [nm]</td>
<td>68</td>
<td>134</td>
<td>115.2</td>
<td>119.6</td>
<td>147.4</td>
<td>151.2</td>
<td>137.8</td>
<td>178.8</td>
</tr>
</tbody>
</table>

Figure 3.A3: Reflection phase variation versus incidence angle. (a) Reflection phase at the center wavelength of 800 nm for TE polarized light obtained from a uniform array of meta-atoms corresponding to the black curve shown in Fig. 3.A2b versus incident angle. (b) Same as a for TM polarization.

Device Fabrication

Conventional nano-fabrication techniques were used to fabricate the folded metasurface hyperspectral imager. A 600-nm-thick layer of \(\alpha\)-Si was deposited on one side of a 1-mm-thick fused silica substrate through a plasma enhanced chemical vapor deposition process at 200°C. Parts of the sample corresponding to the transmissive metasurface were covered using a photolithography mask (AZ-5214E positive resist) and the rest of the sample (corresponding to the reflective metasurfaces) was etched about 205 nm in a mixture of \(\text{SF}_6\) and \(\text{CF}_8\) plasma. The mask was then removed using Acetone. All metasurface patterns were then generated in a \(~300\)-nm-thick layer of ZEP-520A positive electron resist (spun for 1 minute at 5000 rpm) using an EBPG-5200 electron beam lithography system. After development of the resist (in ZED-N50 from Zeon Chemicals), a \(~70\)-nm-thick alumina layer was evaporated on the sample in an electron beam evaporator. After lift-off, this layer was used as a hard mask for dry etching the \(\alpha\)-Si layer. The etching process was again done in two different steps, using a photolithography mask to avoid over-etching the thin (i.e., reflective metasurface) areas. First, all of the sample was etched about 395 nm, then the part corresponding to the reflective metasurfaces was covered using a photolithography mask, and the rest of the sample was etched another 205 nm.
The mask was then removed using Acetone. The alumina layer was also removed in a 1:1 solution of H\textsubscript{2}O\textsubscript{2} and NH\textsubscript{4}OH. A \~2-\textmu m-thick layer of SU-8 2002 polymer was spin-coated, hard-baked and cured on the sample to protect the metasurfaces. The output aperture (which was on the same side as the metasurfaces) was defined using photolithography (AZ-5214E positive resist, MicroChemicals) and lift-off. A \~100-nm-thick gold layer was deposited as the reflective surface. To protect the gold reflector, a second layer of SU-8 2002 was used. To define the input aperture, a \~2-\textmu m-thick layer of SU-8 2002 polymer was spin-coated and cured on the second side of the wafer to improve adhesion with gold. The input aperture was then defined in a process similar to the output aperture.

Device characterization procedure

The measurement setups used to characterize the folded hyperspectral imager are schematically shown in Fig. 3.A4. Light from a tunable Ti-sapphire laser (SolsTiS, M-Squared) was coupled to a single mode optical fiber and collimated using a fiber collimator (F240FC-B, Thorlabs). A fiber polarization controller and a free space polarizer (LPVIS100-MP2, Thorlabs) were used to control the input light polarization and neutral density filters were used to control the light intensity. The hyperspectral imager was placed at the center of a rotating stage where the collimator is mounted. The illumination angle on the input aperture of the hyperspectral imager was varied by the rotating stage. The focal plane of the hyperspectral imager, located \~1 mm away from the output aperture, was then imaged using a custom built microscope (objective: 20\times LUCPlanFl, NA=0.45, Olympus; tube lens: AC254-200-C-ML, Thorlabs; camera: CoolSNAP K4, Photometrics). Since the field of view was \~1.36 mm (limited by the \~15-mm image sensor, and the \~11.1\times magnification) while the spread of angles focused on the image plane are larger than field of view (Fig. 3.3), the objective was scanned on the image plane to cover the focal plane at each wavelength and angle as shown in Fig. 3.A5a. The measurements were performed at several wavelengths and angles, some of which are shown in Fig. 3.A5a. The measurements were also performed at a second set of wavelengths and at normal incidence angle (751.5 nm, 801.5 nm, and 851.5 nm). These results are summarized in Fig. 3.A5b. Another set of wavelengths with an angular separation of 0.1° around 0° were also measured [Fig. 3.A5c].

The setup used to measure the focusing efficiency of the hyperspectral imager is similar to the one used in [133]. The efficiencies for TE and TM polarizations at 0° incidence angle versus wavelength were measured independently using a linear
polarizer. As shown in Fig. 3.A6, the efficiency of the hyperspectral imager is almost uniform, with a maximum at the center wavelength of 800 nm. The average efficiency of the hyperspectral imager is measured to be \( \sim 10.2\% \).

To perform hyperspectral imaging for an object, an imaging optics unit is placed in front of the hyperspectral imager input aperture to map the points on one lateral scan of the object into collimated angles spreading within the range of \(-13^\circ\) to \(13^\circ\) incident onto the device [Fig. 3.A4c]. The imaging optics [Fig. 3.A4c] consists of a 10× objective lens (M-Plan Apo 10, Mitutoyo) and a 4-f system, with the focal lengths of 20 cm (LB1945-B-ML, Thorlabs) and 3 cm (AC254-030-B-ML, Thorlabs), that image the back focal plane of the objective onto the device aperture. A supercontinuum laser (Fianium Whitelase Micro, NKT Photonics) was used to illuminate the object. Two filters, one short-pass (840 nm SP) and the other one long-pass (750 nm LP), were used to filter out of band light. The hyperspectral sample was then scanned by a picomotor (8302, Newport) connected to the stage. At each step in the scan, the spectral data for different angles (-15 to +15 degrees) in the 750–840 nm were captured. The rate of capturing is 4 \( \mu \)m/scan (400 steps/seconds, where each step was about 10 nm and each scan takes 1 second), such that the whole sample was scanned in about 250 line-scans.

The setup shown in Fig. 3.A4d was used to measure the spectral content of the object for comparison with the data obtained from the metasurface hyperspectral imager. The Ti-sapphire tunable laser used in Fig. 3.A4d is tuned with high precision to the required wavelengths and the transmission intensity of the object of interest, is calculated. However, the high spatial coherence of the laser results in formation of speckle patterns. To reduce this spatial coherence, a custom-made rotating diffuser was used in the setup. To obtain the transmission intensity of the object, the tunable laser is swept from 750 nm to 840 nm with increments of 2 nm. At each step the transmitted intensity through the object is captured and subtracted from the "base intensity", measured with the object removed from the path. The results of measurements using the two techniques [Fig. 3.A4c and 3.A4d] are shown in Supplementary Video 1. The optimal expected operation of the designed sample is also shown in Supplementary Video 2. As seen in the videos, by increasing the input wavelengths from 750 nm to 840 nm, the minimum transmission line moves upwards.

**Reconstruction of the hyperspectral data-cube**

By extracting the angles and wavelengths of the scans (as shown in Fig. 3.A7a) using
Figure 3.A4: Measurement setups used for characterization, efficiency calculation, and hyperspectral imaging. (a) Schematic of the setup used for characterization of the folded metasurface hyperspectral imager. (b) Experimental setup used to measure the efficiency of the hyperspectral imager. (c) Measurement setup used to obtain the hyperspectral data-cube of the designed object using the metasurface hyperspectral imager. (d) Experimental setup used to measure the hyperspectral content of the object using a tunable laser to verify the results.
Figure 3.A5: Characterization results of the metasurface hyperspectral imager. (a) Measured intensity profiles at several wavelengths and angles. (b) Measured intensity profiles of two sets of 0-degree incident beams at wavelengths separated by 1.5 nm. The inset shows 2D intensity profiles, demonstrating two resolvable peaks. (c) Measured 2D intensity profiles for sets of two beams separated by 0.1° around 0°. 1D cuts showing two resolvable peaks are plotted on the right. Scale bars: 10 μm for all parts.

Figure 3.A6: Measured efficiencies for TE and TM polarizations. Focusing efficiency of the hyperspectral imager versus wavelength, measured for both TE and TM polarizations. The symbols illustrate the measured data and the solid lines are eye guides.
the data in Fig. 3.3b of the main manuscript, and applying filters at angles $|\theta| > 13^\circ$ and wavelengths larger than 840 nm, each captured frame of the line-scan is converted to a cross section of the hyperspectral data-cube. By stacking these cross-sections for all line-scans as shown in Fig. 3.A7b the data-cube is reconstructed.

Figure 3.A7: Reconstruction of the hyperspectral data-cube. (a) Measured intensity on the image plane corresponding to the first line scan. The black points show the corresponding map of wavelengths (750–840 nm) and angles (-13°–13°). (b) The stack of scanned frames showing how the object data-cube is reconstructed.

Hyperspectral sample design and fabrication

To design a sample with noticeable transmittance spectrum variation over its area, we used the metasurface unit cell shown in Fig. 3.A8a. The width of $\alpha$-Si nano-posts in the unit cell was swept from 164 nm to 220 nm, resulting in the transmission minimum moving from 750 nm to around 850 nm, as shown in Fig. 3.A8b. The lattice constant and nano-post height were kept constant at 300 nm and 480 nm, respectively, to avoid unwanted diffraction and facilitate fabrications. Sweeping of the resonance wavelength according to Fig. 3.A8b enables the design and realization of samples with spatial variations in their transmission spectrum. As an example, we designed and fabricated a "Caltech Logo" sample, shown in Fig. 3.A8c. The inset of Fig. 3.A8c shows scanning electron micrographs of the fabricated sample. As shown in Fig. 3.A8c, the transmission dip wavelength changes from 850 nm in the top to 750 nm in the bottom. They sample was fabricated in a process similar to that used for making the metasurface hyperspectral imager itself. Optical images of the sample at two different fabrication steps, after lift-off and after etching, are shown in Fig. 3.A8c.
Figure 3.A8: Design and fabrication of the hyperspectral test sample. (a) Unit cell of the hyperspectral sample. (b) Transmission spectra of a periodic array of the unit cells shown in a for side length swept from 164 nm to 220 nm. (c) Desired transmission dip wavelength of the fabricated sample, next to scanning electron micrographs of part of the sample (scale bars: 100 μm and 2 μm for left and right images). Schematic of two fabrication steps, after lift-off and after etch, alongside optical microscope images of the sample at these two steps.

3.4  Appendix: Additional information and discussion for random phase mask metasurface hyperspectral imager

Reconstruction using decorrelation of speckle patterns

By incoherent illumination of the sample object, the total intensity on the image plane is written as the summation of the intensities of different wavelengths:

\[ I_{tot} = \sum_{i=1}^{N} I_{\lambda_i} = \sum_{i=1}^{N} O_{\lambda_i} * PSF_{\lambda_i} \]  \hspace{1cm} (3.A1)

Assuming a total decorrelation of point spread functions at different wavelengths \((PSF_{\lambda_i} * PSF_{\lambda_j} = \delta_{ij})\), and taking the Fourier transform of equation 3.A1, we reach:

\[ \sum_{i \neq j}^{N} FFT\{I_{\lambda_i}\}FFT\{I_{\lambda_j}\}^c \approx 0 \]  \hspace{1cm} (3.A2)

\[ \Rightarrow FFT\{I_{tot}\}FFT\{PSF_{\lambda}\}^c \approx FFT\{I_{\lambda_j}\}FFT\{PSF_{\lambda_j}\}^c \]
Therefore each $O_{\lambda_j}$ can be calculated as:

$$O_{\lambda_j} = \text{FFT}^{-1} \left[ \frac{\text{FFT}(I_{\lambda_j}) \text{FFT}(PSF_{\lambda_j})^c}{|\text{FFT}(PSF_{\lambda_j})|^2} \right] = \text{FFT}^{-1} \left[ \frac{\text{FFT}(I_{\text{tot}}) \text{FFT}(PSF_{\lambda_j})^c}{|\text{FFT}(PSF_{\lambda_j})|^2} \right]$$  \hspace{1cm} (3.A3)

The problem can be formulated as the following:

$$y^{m \times 1} = \sum_{i=1}^{N} A_{\lambda_i}^{m \times n} x_{\lambda_i}^{n \times 1}$$  \hspace{1cm} (3.A4)

where $y^{m \times 1}$ is the measurement vector and $A_{\lambda_i}^{m \times n}$ for $i = 1, \ldots, N$ are measurement matrices corresponding to PSF at each wavelength. In the same way as equation 3.A2 the relation between measurement matrices can be written as:

$$A_{\lambda_i}^T A_{\lambda_i} = \delta_{ij} I$$  \hspace{1cm} (3.A5)

In the actual case the assumption in equation 3.A5 does not hold true as the $\|A_{\lambda_i}^T A_{\lambda_i}\|_2 \approx 0.002$. Therefore for the case of N=300 a considerable error in reconstruction of each wavelength is caused and the reconstruction approach based on deconvolution would not be effective as shown in Fig. 3.A9 unless further requirements about the object such as sparsity and/or spatial spectral separation is supposed.

![Reconstruction at different wavelengths](image)

**Figure 3.A9: Object reconstruction at different wavelengths by deconvolution.**
The original and reconstructed objects using deconvolution by the point spread function at each wavelength as the number of wavelength increase. The size of the data-cube is $(51 \times 51 \times 21)$. The width of each image is 100 $\mu$m.

**Compressive sensing reconstruction using joint sparse and low-rank matrix assumption**

The optimization problem using joint sparse and low rank representation matrix can be defined as:

$$\hat{x} = \text{argmin} \left( \|Ax - b\|_2^2 \right)$$

s.t. $\|x\|_1 < \epsilon_1$, $\|y\|_* < \epsilon_2$
Where $A$ is the measurement matrix, $b$ is the observation matrix and $\gamma$ is the operator to obtain the matrix representation of the data cube from its spectral bands. The strict constraints in equation 3.4 can be moved to the objective function by using the Lagrangian to penalize big values of $\ell_1$ and $\ell_*$. The more relaxed version of the problem can be formulated as:

$$
\hat{x} = \arg\min_{x} \left( \|Ax - b\|^2_2 + \tau \|x\|_1 + \mu \|y\|_* \right)
$$

$$
\tau, \mu > 0
$$
Many potential applications of metasurfaces have been investigated so far such as two-photon microscopy [219], retro-reflection [131], image differentiation [220] and many others as discussed in chapter 1 of this dissertation [221]. In this chapter other not-yet-fully-explored applications of optical metasurfaces and the implementation of optimization techniques to realize more efficient optical metasurfaces will be discussed.

4.1 Inverse design of high efficiency optical metasurfaces

In many instances, optical metasurfaces require to be optimized to perform their intended functionalities efficiently. Optimization techniques are essential to be employed to achieve this goal. In most of the optimization techniques an objective function is defined to be either minimized or maximized based on the existing degrees of freedom in the problem. Suppose there are N degrees of freedom in a convex optimization problem. To achieve the optimization goal, the derivative of the objective function with respect to all degrees of freedoms needs to be calculated and the parameters need to be adjusted accordingly in order to maximize or minimize the objective function in that direction. Thus, N forward computations are required to calculate the derivative. In a large optimization problem as the one used for the design of optical metasurfaces the degrees of freedom are in the order of millions, so the conventional way to calculate the gradient is inefficient and/or computationally heavy.

Adjoint optimization is a versatile and powerful technique, used in the design of optical and electromagnetics devices, such as couplers, beam splitters [222], integrated photonics components [223, 224], and photonic crystals [225]. The technique reduces the number of calculations from N to 2 [226–228]; one for the forward problem and the other one for the inverse or adjoint problem. Different utilizations of the adjoint method in designing optical metasurfaces for various applications either in single layer [125, 229, 230] and or multi-layer metasurfaces [137, 231–233] have been reported. One of the areas where the adjoint method can be employed is the design of efficient metasurface lenses over a broader range of spectrum rather than the conventional design which is performed at single wavelength. Here, we present a
technique where the adjoint optimization is used to realize a high efficiency focusing lens over \( \sim 200 \text{ nm} \) bandwidth in near-Infrared (750 nm to 950 nm).

Figure 4.1 shows the transmission intensity and phase of a unit cell made of \( \alpha \)-Si nanoposts where the height and lattice constant are 760 nm and 275 nm respectively. The transmission phase and intensity are plotted for three wavelengths of 750 nm, 850 nm and 950 nm. The height of nanoposts (760 nm) is selected in such a way that for all wavelengths we achieve at least \( 2\pi \) phase as the size of nano posts are varied from 60 nm to 200 nm.

![Diagram of unit cell](image)

**Figure 4.1:** Transmission intensity and phase of the unit cell for different wavelengths. (a) Unit cell used for the design of an efficient optical metasurface. (b) The transmission intensity and phase plotted vs widths of \( \alpha \)-Si nanopost at three wavelengths of 750 nm, 850 nm and 950 nm.

The conventional method for designing a spherical aberration-free lens starts by implementing the phase profile at the center wavelength (850 nm). By doing so, the phase profile at the center wavelength would be similar to the one for a spherical aberration-free lens. Here we design a lens with the diameter of 100 \( \mu \text{m} \) and focal length of 250 \( \mu \text{m} \). Figure 4.2a shows the phase profile of spechrical aberration-free lens designed at the center wavelength of 850 nm along with the post dimensions associated with the phase profile extracted from the curve in Fig. 4.1b. At wavelengths far 850 nm, especially those at the edge of the spectral band, a different phase profile is expected, resulting in poor focusing efficiencies. The effect is shown in Fig. 4.2b and c where the focal plane and axial plane intensity distributions are plotted. Specifically, the intensity concentration at 750 nm has occurred at another focal plane close to 150 \( \mu \text{m} \), shown in Fig. 4.2c, leading to a considerable reduction of focusing efficiency at the desired image plane.

To enhance the focusing efficiency of the device at other wavelengths and improve its performance an optimization problem is designed as the following: the objective
Figure 4.2: Focusing performance of a spherical aberration-free lens designed at 850 nm. (a) Phase profile and the associated post dimensions for the lens designed at 850 nm. The lens has a diameter of 100 μm and focal length of 250 μm. (b) Focal plane intensity plotted for three wavelengths of 750 nm, 850 nm and 950 nm. (c) Axial plane intensity distribution shown at three wavelengths of 750 nm, 850 nm and 950 nm.

The implementation of an adjoint optimization scheme for increasing the focusing...
4.2 Optical vortex coronographs

Introduction

Optical vortex coronographs are a type of optical instruments used for imaging very faint objects near very bright ones, such as the exoplanets located close to their bright host star. Compared to other type of coronagraphs, vortex coronagraphs (VCs) are an attractive solution for future earth (ground-based) [234–236] or space [237] telescopes due to their high throughput, large spectral bandwidth and low sensitivity to low-order aberrations [238].

For the first time a method to observe dim exoplanets that eliminates light from the parent star across the entire exit pupil without sacrificing light from the planet was shown in [239]. Up until recently, telescopes were only able to image exoplanets which were quite large, hot and intensely radiative and separated widely from their...
In 2010, the detection of all three HR.8799 planets using a small-aperture 1.5 m system was reported [240]. So, well-corrected coronographs are able to detect exoplanets even closer to their host stars. VCs also allow for further reduction in the size of prospective space telescopes designed to image very faint terrestrial planets. The next generation of VCs are aimed at dimming multiple sources of light to be used for imaging planets around multi-star systems [241].

Two realizations for VCs are available and widely used today: vectorial and scalar. The distinction between the two is based on how they turn the starlight into an optical vortex. Vectorial VCs use half wave plates with spatially varying fast axis, which results in polarization dependent geometric phase shift [238]. Vector phase masks are usually fabricated using single layer or multiple layers of patterned liquid crystals [242–245], subwavelengths gratings [246], photonic crystals [247], metamaterials [248] and/or rotationally symmetric prisms [249].

Here, we study the realization of vortex coronagraphs by means of optical metasurfaces. Specifically, we present the simulation, design and fabrication of a planar VC with charge 6, operating at 700 nm and having a diameter of 5 mm. One of the advantages of a metasurface VC over other types of coronograph is the polarization independent phase shift. Furthermore, the designed VC is required to have the minimal retardance error across its entire bandwidth of 10%. Figure 4.4 shows the schematic for a vortex of charge 6 simulated at three wavelengths of 665 nm, 700 nm and 735 nm. As illustrated in Fig. 4.4b the intensity at the center of the image plane is $\sim 10^{-5}$ of its surrounding ring of intensities.

**Simulation and design**

The transmissive unit cell used for the design of the planar vortex is shown in Fig. 4.5a. Using the RCWA technique [171] the transmission coefficients for a periodic array of unit cells were obtained. For each specific set of nano-post dimensions, a uniform array of the $\alpha$-Si nano-posts was illuminated with a plane wave at the wavelength of 700 nm under normal illumination angle, and the transmitted amplitudes and phases were extracted. The nano-posts are symmetric and have the same widths ($W$) in $x$ and $y$ directions. To choose the height of nano-posts, we performed the simulations for nano-posts of square cross-sections with different heights and side lengths [Fig. 4.5b-d]. The height of nano-posts was then chosen to minimize the variation of the derivative of the phase with respect to wavelength for different side lengths, while
Figure 4.4: Simulation of a vortex coronagraph with charge 6. (a) Phase profile a VC of charge 6 having a diameter of 5 mm. (b) Simulated intensity and phase on the image plane located 100 mm away from the vortex at three wavelengths of 665 nm, 700 nm and 735 nm.

providing a full $2\pi$ phase coverage and high transmission. In this way the retardance error across the entire bandwidth of 10% is less than 5%. The value can be further reduced by using other dispersion engineering techniques available in the design of optical metasurfaces [250].

The post height and lattice constant were chosen to be 300 nm and 430 nm, respectively. These two values will achieve more than full $2\pi$ phase coverage while minimizing variation of the reflection phase derivative across the band, as the side lengths of nano-post are swept from 60 nm to 200 nm. [Fig. 4.5e].

Device Fabrication

The vortex coronagraph was fabricated using the standard micro- and nano-fabrication techniques. The fabrication process started with the deposition of a 430-nm thick layer of $\alpha$-Si on a 500-μm-thick fused silica substrate. The metasurface pattern was defined in a ~ 300-nm-thick ZEP-520A positive electron-beam resist (behaving as a soft mask) using electron-beam lithography. The bright field and dark field optical microscope image of the sample after development is shown in Fig. 4.5f. After development of the resist, a ~70-nm-thick layer of aluminum oxide ($Al_2O_3$) was deposited on the sample using electron beam evaporation and lifted off to invert the pattern. The optical microscope image of the sample after lift-off is demonstrated in Fig. 4.5f. The aluminum oxide is then used as a hard mask in the reactive ion etching
Figure 4.5: Unit cell, design curve and images of the fabricated vortex. (a) Unit cell used to implement the metasurface vortex. (b) Phase of nano-posts versus their side lengths and heights. (c) Transmission of nano-posts versus their side lengths and heights. (d) Phase derivative of square cross-section nano-posts with respect to wavelength plotted versus their side length and height. The black line shows the height (430 nm) where the phase derivative is minimal. (e) The dictionary which shows the selected side length of nano-post at each phase value. (f) Optical image of the fabricated vortex at different fabrication steps (right) next to scanning electron micrographs of part of the sample (right).

of the $\alpha$-Si layer. Finally, the aluminum oxide mask is removed in a solution of hydrogen peroxide and ammonium hydroxide with 1:1 ratio. The optical microscope image of this step with a scanning electron micrograph of the central part of the fabricated vortex are shown in Fig. 4.5f.

4.3 Appendix: Additional information and discussion for inverse design of high efficiency optical metasurfaces

The formulation of adjoint optimization problem for maximizing the efficiency across different wavelengths is as following:

The metasurface transmission mask is defined as $t(D)$ which is the function of post-dimensions. The propagation operator from the metasurface to the image plane is $A$ which is independent of device parameters, the input and output fields are defined as $E_{in}$ and $E_{out}$ respectively. Therefore, $E_{out} = A \cdot E_{in}$. The optimization goal is defined as maximizing the intensity at the image plane, which we call it $F$.  

here. So, at each wavelength the objective function is defined as the following:

\[ F = |E_{out}|^2_{x=0, y=0} = E_{out}E_{out}^*_{x=0, y=0} \] \hspace{1cm} (4.A1)

The derivative of objective function with respect to the post dimensions is calculated as:

\[
\frac{\partial F}{\partial D} = \frac{\partial F}{\partial E_{out}} \frac{\partial E_{out}}{\partial D} + \frac{\partial F}{\partial E_{out}^*} \frac{\partial E_{out}^*}{\partial D} \\
= E_{out}^*A \frac{\partial t}{\partial D} E_{in} + E_{out}A^* \frac{\partial t^*}{\partial D} E_{in}^* \\
= 2 \Re \left( E_{out}^*A \frac{\partial t}{\partial D} E_{in} \right) \] \hspace{1cm} (4.A2)

where the adjoint problem is defined as:

\[ E_{out}^*A = \left( A^\dagger E_{out}^* \right)^\dagger \] \hspace{1cm} (4.A3)

and \( A^\dagger \) is the back-propagation operator acting on \( E_{out}^* \), as the point source on the image plane. The total objective function gradient is calculated as the weighted summation of the gradients for the objective functions at each wavelengths.

\[ F_{tot} = \sum_{i=1}^{N} \omega_i F_i \rightarrow \nabla F_{tot} = \frac{\partial F_{tot}}{\partial D} = \sum_{i=1}^{N} \omega_i \frac{\partial F_i}{\partial D} \] \hspace{1cm} (4.A4)

where \( \omega_i = \frac{2}{N} - \frac{F_i}{\sum_i F_i} \)

Using the gradient decent method the post dimensions \( (D) \) is updated at each step as the following:

\[ D_{n+1} = D_n + \gamma \nabla F_{tot} \] \hspace{1cm} (4.A5)
CONCLUSION AND OUTLOOK

The field of optical metasurfaces has been growing rapidly in the past few years. Optical metasurfaces have brought unique capabilities in increasing the resolution and efficiency of optical systems, while being integrated and maintaining the planar form. In addition, they were used to demonstrate novel optical functionalities, not directly achievable using conventional optics. Furthermore, optical metasurfaces exhibit advantages over conventionally-designed bulky optical elements, such as the sub-wavelength thickness, planar form factor, compatibility with standard micro- and nano-fabrication processes, potential low-cost batch fabrication, the capability in controlling different degrees of freedom of light and many others which have brought great attention to them for realization of the next-generation compact and high-performance optical systems.

One of the key promises of optical metasurfaces is the development of compact and planar optical systems. On the other hand, implementation of thin optical systems requires reducing both the size of the optical element and the spacing between them, which relates to the requirement for free-space propagation in many optical systems. Although the former has been shown in realization of optical devices, the later has been uninvestigated for many years. The folded metasurface optics platform introduced in this dissertation, significantly reduces the total size of optical systems by using the substrate as the propagation medium and folding the optical paths so they have partial overlap. As proof of concept, two practical optical systems, i.e. a compact folded metasurface spectrometer and a miniaturized metasurface hyperspectral imager were demonstrated. The potential applications of the proposed platform is expected to go beyond what was explored here, including on-chip interferometers, imaging systems and other advanced optical and/or optoelectronic systems. All-optical neural networks [251, 252] is the other potential area we envision for optical metasurfaces, where they can be employed as layers, and specific nonlinear configurations such as saturable absorbers, electromagnetically induced transparency (EITs) can act as the nonlinear activation functions [253, 254].

Despite all of the advances in the field of metasurfaces in the past few years, several fundamental and practical challenges have still remained unsolved. From a
fundamental perspective, the first challenge is the actual number of available degrees of freedom in a specific volume of a metasurface device, which is proportional to the number of distinguished functionalities the device can achieve. This value becomes more important in the design of multi-functional metasurfaces, where optimization techniques are used to further improve their performance [125, 230, 255] but the maximum number of independent functionalities are still unknown. The second challenge is the limitation on the realization of achromatic optical metasurfaces for imaging. This limitation results in limited operation bandwidths, small sizes and low numerical aperture devices. None of the current dispersion-engineered optical metasurfaces are able to achieve simultaneously large aperture size (~ few millimeters) and high numerical apertures. The other challenge is the existence of locality, angular memory effect and weak couplings between meta-atoms, all of which pose problems for devices working under large deflection angle constraints. Therefore, a new design methodology to take into account all of these and enable large-area non-periodic structures with high efficiency is highly required.

From a practical perspective, one of the challenges in realization of dielectric metasurfaces in the visible range is the unavailability of a low-loss high-index material to provide low coupling between meta-atoms and also full $2\pi$ phase coverage. Therefore, many implementations of optical metasurfaces in the visible [118, 256–258] are still lacking high-efficiency compared to those fabricated in the near-IR. The other practical challenge is the mass production of optical metasurfaces in large-scale. This requires a manufacturing process compatible with existing low-cost foundry technologies. Large-scale fabrication processes like DUV lithography [259], roll-to-roll nano-imprint [260], and soft lithography have been applied to address the current challenge, however many steps still need to be taken to make this practically achievable. Another category of optical devices which are highly desirable for applications like beam steering, spatial light modulations, dynamic holographic displays and many others are tunable metasurfaces. Despite the great effort in demonstration of several tunable metasurfaces [261–263], none of them still compete with the commercially available liquid crystal based SLMs, in terms of tuning-speed, efficiency and resolution.

With all the above facts in mind, optical metasurfaces are still assumed as a potential candidate for compact realization of optical elements and systems designed for various narrow-bandwidth applications including wavefront engineering, cloaking, selective absorbers, imaging and leaky-wave radiation.
BIBLIOGRAPHY


84


