Boston University

OpenBU **http://open.bu.edu**

Boston University Theses & Dissertations Boston University Theses & Dissertations

2024

Directional photodetectors based on plasmonic metasurfaces for advanced imaging capabilities

https://hdl.handle.net/2144/48870

Downloaded from DSpace Repository, DSpace Institution's institutional repository

BOSTON UNIVERSITY

COLLEGE OF ENGINEERING

Dissertation

DIRECTIONAL PHOTODETECTORS BASED ON PLASMONIC METASURFACES FOR ADVANCED IMAGING CAPABILITIES

by

JIANING LIU

B.S., Harbin Institute of Technology, 2018
M.S., Boston University, 2022

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2024

© 2024 by
JIANING LIU
All rights reserved

Approved by

First Reader ____
Roberto Paiella, Ph.D.
Professor of Electrical and Computer Engineering
Professor of Materials Science and Engineering

Second Reader ____
Lei Tian, Ph.D.
Assistant Professor of Electrical and Computer Engineering
Assistant Professor of Biomedical Engineering

Third Reader ____
Abdoulaye Ndao, Ph.D.
Assistant Professor of Electrical and Computer Engineering
University of California, San Diego

Fourth Reader ____
Alexander Sergienko, Ph.D.
Professor of Electrical and Computer Engineering
Professor of Physics

ACKNOWLEDGMENTS

How time flies, the scene of stepping in the office and lab at Boston University still shows up clearly in my mind. These seven years of research, study, and life, cannot be simply concentrated into one book. So many people around me offered invaluable support throughout my journey of completing my Ph.D. degree.

To my beloved family, my father Shenghui Liu, and mother Xiaoning Wang. Everything started with you. I appreciate your open-mindedness in supporting my participation in the international exchange program for undergraduate research. This enabled me to successfully match with a supervisor and promptly engage in research projects. Despite being far away, you've always been attentive, providing emotional warmth and generous financial support. My father, you're the best doctor in the world. During the COVID-19 pandemic, you not only guided me on prevention but also resolved my allergy that troubled me for two months. You are continuously offering advice on healthy living and by example. My mother, you are the epitome of caring. You perceive my subtle needs, discussing future plans and current difficulties with me in a gentle and egalitarian manner, strengthening my soul. Although you may not comprehend the specifics of my research, you've paved a solid path for my advancement.

I am grateful for my supervisor, Roberto Paiella. Your meticulous cultivation of me and your personal practice in research have proven that you are an excellent research mentor. Over the past seven years, your precise judgment and extensive knowledge have been the inspiration for me to remain humble and strive for

progress. Also thanks to my committee members who spent time and effort to contribute advice on my research.

To my role models, senior peers Leonard Kogos, Yuyu Li, and Xiaowei Wang. With endless patience, you have taught me many skills step by step and have often clarified my doubts. Also, my gratitude goes to my fellow lab mates, acquaintances in the field of ECE, as well as my wonderful friends who formed a supportive community and enriched my life. I especially thank my lovely ex-girlfriend who accompanied me for 9-month that became my most brilliant time in both research and life.

Finally, I would like to express my gratitude towards the organizations that supported this work. Many challenging fabrication tasks were conducted at the Center for Nanoscale Systems (CNS) of Harvard and Brookhaven National Laboratory, where staff and students are very professional and friendly.

Sincerely,

Jianing Liu

DIRECTIONAL PHOTODETECTORS BASED ON PLASMONIC METASURFACES FOR ADVANCED IMAGING CAPABILITIES

JIANING LIU

Boston University College of Engineering, 2024

Major Professor: Roberto Paiella, Ph.D., Professor of Electrical and Computer Engineering, Professor of Materials Science and Engineering

ABSTRACT

With the continuous advancement of imaging technologies, imaging devices are no longer limited to the exclusive measurement of optical intensity (at the expense of all other degrees of freedom of the incident light) in a standard single-aperture configuration. Increasingly demanding applications are currently driving the exploration of more complex imaging capabilities, such as phase contrast imaging, wave front sensing, optical spatial filtering, and compound-eye vision. Many of these applications also require highly integrated, lightweight, and compact designs without sacrificing performance. Thanks to recent developments in micro- and nanophotonics, planar devices such as metasurfaces have emerged as a powerful new paradigm to construct optical elements with extreme miniaturization and high design flexibility. Sophisticated simulation tools and high-resolution fabrication techniques have also become available to enable the implementation of these compact subwavelength structures in academic and industrial labs. In this dissertation, I will present my work aimed at achieving directional light sensing by directly integrating composite plasmonic metasurfaces on the illumination windows of standard planar

photodetectors. The devices developed in this work feature sharp detection peaks in their angular response with three different types of behaviors: symmetric around the device surface normal, asymmetric with nearly linear angular variations around normal incidence, and geometrically tunable single peaks up to over 60 degrees. The performance of the proposed metasurfaces has been optimized by full-wave numerical simulations, and experimental devices have been fabricated and tested with a custom-designed measurement setup. The measured angular characteristics were then used to computationally demonstrate incoherent edge enhancement for computer vision and quantitative phase-contrast imaging for biomedical microscopy. Importantly, the device fabrication process has also been upgraded to wafer scale, further promoting the possibility of batch-production of our devices.

TABLE OF CONTENTS

ACKNOWLEDGMENTS iv

ABSTRACT vi

TABLE OF CONTENTS viii

LIST OF FIGURES xi

LIST OF ABBREVIATIONS xiii

1 Intruduction 1

2 Plasmonic Directional Image Sensors for Optical Spatial Filtering 6

2.1 Background 6

2.2 Introduction 8

2.3 Design and Simulations 11

2.4 Device Fabrication and Measurement 15

2.5 Coherent Transfer Function and Phase Contrast Imaging 20

2.6 Optical Transfer Function and Optical Spatial Filtering 29

2.7 Conclusion 37

3 Asymmetric Directional Photodetectors for Quantitative Phase Contrast Imaging 38

3.1 Introduction 38

3.1.1 Background 38

3.1.2 Principle of Design and Functionality 40

3.2 Design and Simulations 45

3.3 Fabrication and Measurement results 49

3.4 DPC Imaging 51

3.4.1 Physical Model and DPC Imaging Simulation 51

3.4.2 Minimum Detectable Phase Contrast 54

3.4.3 Single-Shot Quantitative Phase Reconstruction 57

3.5 Conclusion.. 63

4 Gradient-Metasurface Directional Photodetectors.. 63

4.1 Introduction.. 64

4.3 Design and Simulations 66

4.3 Experimental Results.. 71

4.4 Conclusion.. 75

5 Wafer-Scale Upgrade of Sample Fabrication Process.. 75

5.1 Purpose of the Upgrade 76

5.2 Sample Fabrication Process 80

5.2.1 Wafer Preparation... 80

5.2.2 Process Flow... 82

5.2.3 Recipes for Important Steps.. 89

5.2.4 Photomasks Preparation 95

5.3 Probe Station 97

5.2.1 Introduction to the Design 97

5.2.2 Standard Operation Procedure 99

Conclusion..105

A CTF Calculation 105

B OTF Calculation 111

B.1 Derivation of the Filter OTF 111

B.2 Estimation of the Transverse Coherence Length 113

B.3 Calculation of the Differential OTF 114

B.4 MATLAB Code for OTF 116

C Code for Phase Contrast Imaging Simulations 123

D Derivation and Code for Quantitative DPC Imaging 131

E Failures in Fabrications 137

E.1 Outcome of Moisture in SiO_2 Layers 137

E.2 Attempt of Using HSQ as the Insulation Layer 138

E.3 Attempt of Making the Slits with EBL 139

E.4 Sputtering to Make SiO_2 Layers 141

BIBLIOGRAPHY 142

CURRICULUM VITAE 148

LIST OF FIGURES

2.1 Plasmonic directional image sensors.. 12

2.2 3D schematic of device structure.. 15

2.3 Illustration of device measurement setup with major optical components.......... 18

2.4 Schematic illustration of the physical model used to evaluate the CTF contribution $t_+(k)$ from the slits on the right of the NP array.. 20

2.5 Coherent transfer function of the plasmonic directional image sensors of Figure 2.1. 22

2.6 Phase imaging results... 27

2.7 Optical transfer function of the plasmonic directional image sensors of Figure 2.1... 31

2.8 Incoherent edge detection protocol 33

2.9 Incoherent edge detection results 36

3.1 Phase contrast imaging with angle-sensitive photodetectors... 40

3.2 Asymmetric metasurface photodetectors 42

3.3 Performance of asymmetric metasurface photodetectors 47

3.4 Computational phase contrast imaging results 52

3.5 Minimum detectable phase contrast analysis.. 55

3.6 Computational quantitative phase imaging results 58

3.7 Additional phase contrast imaging results 62

4.1 GMS design 67

4.2 SEM images 71

4.3 Measurement results for a GMS device based on the design of Figure 4.1 72

4.4 Geometrical tunability of the angle of peak detection θ_p 74

5.1 A photomask design with photolithography patterns on different positions 77

5.2 EBL patterning work flow 78

5.3 Structure of a completed device 82

5.4 SEM images of a metasurface patterned by EBL. 84

5.5 Photography of a wafer after the 7 steps before dicing 85

5.6 FIB patterning window 87

5.7 Photomask designs 95

5.8 Front view of the probe station 98

5.9 Workflow of focusing on the tips 100

A.1 Schematic illustration of the physical model used to evaluate the CTF 105

A.2 FDTD simulation results..110

E.1 Photography of a sample with metal films damaged due to inflation of water vapor...137

LIST OF ABBREVIATIONS

AFM	Atomic-Force Microscope
BOE	Buffered Oxide Etch
CMOS	Complementary Metal Oxide Semiconductor
CNN	Convolutional Neural Network
CTF	Coherent Transfer Function
DC	Direct Current
DI	De-Ionized
DPC	Differential Phase Contrast
EBL	Electron-Beam Lithography
FDTD	Finite-Difference Time Domain
FIB	Focused-Ion Beam
FOV	Field Of View
FWHM	Full Width Half Maximum
GMS	Gradient Metasurface(s)
GSP	Gap Surface Plasmon
HSQ	Hydrogen Silsesquioxane
IPA	Isopropanol
MIBK	Methyl Isobutyl Ketone
NPGS	Nanometer Pattern Generation System
OTF	Optical Transfer Function
PECVD	Plasma-Enhanced Chemical-Vapor Deposition
PMMA	Polymethyl Methacrylate
RIE	Reactive-Ion Etching
RMS	Root Mean-Square
SEM	Scanning Electron Microscope
SNR	Signal-to-Noise Ratio
SPP	Surface Plasmon Polaritons

CHAPTER ONE

Introduction

As is seen in daily life, traditional imaging devices embrace a configuration of a single-aperture system with lenses and an array of imaging sensors, an analogy to human eyes. In such imaging systems, the lateral distribution of intensity of the incident light is projected by the lens system and captured at the optical sensor, within a limited field of view. However, there are two limitations that hinder a wider realm of applications. First, all other information of the incident light is lost, including the local propagation direction and polarization. Second, there is a tradeoff between the field of view and spatial resolution. In more advanced applications in the modern era, these limitations are to be broken through. For example, to visualize transparent biological cells, we ask for phase imaging capability. In product quality control such as wafer surface flatness, or to obtain 3-dimensional light-field, wave front sensing is required. For navigation and surveillance tasks, we need motion-sensitive compound-eye cameras where each pixel is a directional sensor. However, existing solutions often result in either bulky or complicated structures, which hinders the realization in many tasks. (Paganin and Nugent, 1998), (Mehta and Sheppard, 2009), (Lam, 2015), (Duparré et al., 2005)

A newly emerging concept of metasurfaces provided a solution of constructing optical elements with unprecedented freedom of designs, meanwhile achieving extreme miniaturization. Metasurfaces are a type of planar structures composing of subwavelength unit cells, and the response to the incident light by each unit cell (such

as transmitted or reflected phase, amplitude, and polarization) mostly depends on its geometry, which can be designed individually according to its role. In another perspective, metasurfaces are a 2-dimensional version of metamaterials. When carefully designed unit cells are assembled into an array in a plane, many optical elements can be achieved, such as focusing lenses, holograms, perfect absorbers, waveguide couplers, and structured beam generators. (Yu and Capasso, 2014), (Mueller et al., 2017), (Li et al., 2017), (Bhattarai et al., 2017)

However, in the existing research, the community mainly focused on optical elements in the free space, to manipulate the propagation of light. In this work, we explore the methodology of implementing unusual response to the incident light directly on traditional optical sensors, by designing metasurfaces that are integrated on top of the active substrate. In the past, a platform to enable direct integration of plasmonic metasurfaces on traditional sensors has been developed by L. C. Kogos, et al (Kogos et al., 2020). It is a metal-dielectric-metal (MDM) sandwich structure where the top metallic nano-particles are freely designed to couple the incidence light into SPPs supported by the bottom metallic film. The dielectric layer in the middle serves as the role of generating gap-surface-plasmons (Pors et al., 2014) which is crucial to the realization of 2π coverage of the metasurface's phase range. Then, according to the principle of extraordinary transmission of subwavelength apertures in metallic films (Ebbesen et al., 1998), a group of slits perforating the MDM layers are on one side of each pixel to scatter down the SPPs into the substrate (optical active material). In his work, directional sensing by gradient metasurfaces was achieved in both

numerical simulations and experimental samples. As a proof-of-concept verification, imaging simulations were performed to showcase the functionality of compound-eye vision, by virtually assembling a group of such directional sensors into an array.

The objective of my research is to explore and improve the engineering possibilities of this new type of imaging devices that transmit incident light only at desired angles, which can push forward the miniaturization of optical systems and a higher degree-of-freedom in designing optical detectors. The proposed sensors are based on the same structural platform introduced above, with new designs of composite metasurfaces. Multiple functionalities enabled by these sensors will be tested by imaging simulations, such as optical spatial filtering for edge detection, quantitative phase imaging, and compound-eye vision. In many modern applications of imaging devices, those capabilities are invaluable to the quality, compactness, and cost-efficiency.

In detail, the next chapter introduces a simple metal-dielectric-metal grating applied on the design platform, which enables a symmetric response of sharp transmission peaks, relative to the incidence angle around 0 degree. An array of such sensors can function as optical spatial filters without any external filtering elements, leading to extreme size miniaturization. Furthermore, they offer the distinct capability to perform multiple filtering operations at the same time, through the use of sensor arrays partitioned into blocks of adjacent pixels with different angular responses. To establish the image processing capabilities of these devices, I will present a rigorous theoretical model of their filter transfer function under both

coherent and incoherent illumination. Next, the measured angle-resolved responsivity of prototype devices will be used to demonstrate two examples of relevant functionalities: (1) the visualization of otherwise invisible phase objects and (2) spatial differentiation with incoherent light. These results are significant for a multitude of imaging applications ranging from microscopy in biomedicine to object recognition for computer vision.

Next in Chapter 3, I will describe another design of the directional detector but with reflectors involved that create an asymmetric dependence of responsivity on angle of incidence around the surface normal. The metasurface design, fabrication, and angle-sensitive operation are demonstrated using a simple photoconductive detector platform. The measurement results, combined with computational imaging calculations, are then used to show that a standard camera or microscope based on these metasurface pixels can directly visualize phase objects without any additional optical elements, with state-of-the-art minimum detectable phase contrasts below 10 mrad. Furthermore, the combination of sensors with equal and opposite angular response on the same pixel array can be used to perform quantitative phase imaging in a single shot, with a customized reconstruction algorithm which is also developed in this work. By virtue of its system miniaturization and measurement simplicity, the phase imaging approach enabled by these devices is particularly significant for applications involving space-constrained and portable setups (such as point-of-care imaging and endoscopy) and measurements involving freely moving objects.

Chapter 4 demonstrates the use of plasmonic GMS to achieve directional detection with large FOV. The resulting devices rely on the phase-matched coupling of light incident at select geometrically tunable angles into SPP modes. This approach has smaller footprint and reduced guided-mode radiative losses (and therefore improved spatial resolution and sensitivity) compared to analogous devices based on diffractive coupling. More broadly, these results highlight a promising new application space of flat optics, where gradient metasurfaces are integrated within image sensors to enable unconventional capabilities with enhanced system miniaturization and design flexibility.

Finally in Chapter 5, I will describe an important upgrade to the fabrication process. In the works presented in this Dissertation, all samples were fabricated with the old technique, in a chip-by-chip style. This method certainly offers the best flexibility and is least prone to mistakes and incidents in the process, but is very slow and complicated. With a more mature set of recipes, better equipment, and a new fabrication work flow, the samples now are able to be fabricated in a whole-wafer scale in most steps. This results in a major improvement to the quality, production speed, and consistency among the samples.

CHAPTER TWO

Plasmonic Directional Image Sensors for Optical Spatial Filtering

2.1 Background

Spatial filtering operations, where different frequency components of an image are selectively transmitted or blocked, play a key role in many high-impact applications in microscopy, photography, and computer vision (Gonzales and Woods, 2018). For example, low-pass filtering can be used to suppress high-frequency noise and accentuate slowly varying features. Similarly, edge detection by high-pass filtering allows for image sharpening as well as segmentation, to distill a highly compressed version of the original image that is easier to store, transmit, and process. In fact, the latter idea is at the core of the initial stage of the visual recognition process, where different filtered versions of the original image are produced for subsequent analysis. The same principle is also observed in the first layer of CNNs, which have emerged as the leading algorithmic approach for many demanding applications in visual data processinsg such as image classification and object recognition (Krizhevsky et al., 2017). These filtering operations can be readily implemented in the electronic digital domain – at the expense, however, of substantial power consumption and processing time. As a result, their adoption in many embedded and mobile edge-computing applications remains a significant challenge (e.g., in autonomous vehicles, augmented reality headsets, and robots, where power and bandwidth are highly constrained).

These considerations have created novel opportunities for optical computing solutions, which in fact are currently enjoying a substantial resurgence of interest

(Wetzstein et al., 2020). Photonics intrinsically offers ultrafast processing bandwidths (essentially at the speed of light) and low power consumption (only limited by optical propagation losses). Of particular relevance in the context of spatial filtering are approaches based on Fourier optics (Goodman, 2005), building on the well-known two-lens 4f imaging system. In this setup, the first lens projects the Fourier transform of the object field onto a pupil mask between the two lenses, where different spatial frequency components are multiplied by different transmission coefficients before they are recombined by the second lens to form a filtered image of the object. However, this system suffers from large form factor and strict alignment requirements, which again limit its portability. In recent years, several nanophotonic structures have been investigated as a means to provide similar functionalities (particularly image differentiation for edge detection) with more compact dimensions and enhanced design flexibility (Silva et al., 2014). Specific examples include phase-shifted Bragg reflectors (Bykov et al., 2014), plasmonic filters (Zhu et al., 2017), gradient metasurfaces (Pors et al., 2015), (Davis et al., 2019), (Zhou et al., 2019), (Zhou et al., 2020), (Huo et al., 2020), diffraction gratings (Fang and Ruan, 2018), and photonic crystal slabs (Guo et al., 2018), (Cordaro et al., 2019), (Saba et al., 2018), (Zhou et al., 2020), all designed to introduce a sharp in-plane-wavevector dependence in their transmission coefficient. The use of Fourier optical filters in conjunction with neural networks is also being investigated extensively (Chang et al., 2018), (Colburn et al., 2019), (Muminov and Vuong, 2020).

2.2 Introduction

In this chapter, I will introduce a different approach where optical spatial filtering is achieved (on a pixel-by-pixel basis) with an image sensor array consisting of specially designed directional photodetectors. Specifically, we employ devices coated with a plasmonic metasurface that only allows for the detection of light incident along a small set of directions, whereas light incident along all other directions is reflected. This novel capability has been demonstrated in the aforementioned recent work focused on a different imaging application (Kogos et al., 2020), i.e., planar lensless compound-eye vision with ultrawide field of view. Similar devices can also be used as optical spatial filters, based on the notion that different spatial-frequency components of an illuminated object correspond to plane waves propagating from the object along different directions. Importantly, with this approach the filter transfer function can be tailored through the design of the metasurface, and different metasurfaces (i.e., different filters) can be applied on different adjacent pixels within the same image sensor array. As a result, multiple filtering operations can be performed simultaneously with the same pixel array. Furthermore, this approach does not require any external optical components other than a standard imaging lens, and therefore is particularly convenient in terms of system miniaturization and alignment simplicity.

In the directional image sensors described below, light incident at the target detection angles is selectively detected via resonant coupling to a guided plasmonic mode. As a result, sharp responsivity peaks at geometrically tunable angles are

obtained, which are particularly well suited to engineer a wide range of transfer functions for high-contrast optical spatial filtering. In contrast, other devices previously used for angle-sensitive vision (Duparré et al., 2004) (Gill et at., 2011) (Yi et al., 2021) feature a more gradual angular dependence and/or limited tunability. At the same time, the role of the guided modes in our directional photodetectors complicates the conceptual analogy with standard optical spatial filters. Therefore, in order to substantiate the image processing capabilities of these devices, here we develop a rigorous theoretical model that quantifies and clarifies the nature of their filter transfer function. Next, we combine this model with the experimental angle-resolved responsivity of prototype samples to demonstrate two examples of relevant spatial filtering functionalities.

In the first example, we use devices featuring high-pass filtering characteristics to visualize a transparent phase-only object, which would otherwise be invisible to standard image sensor arrays. Second, we address the task of edge detection of amplitude objects with incoherent (i.e., natural) illumination. It is well established from Fourier optics that, when the incident light is spatially incoherent, high-pass filtering is generally impossible with a single filter (Goodman et al., 2005), which represents a key limitation of optical-domain image processing. A possible solution is to record two low-pass filtered images of the object of interest with different cutoff frequencies, and then compute their difference (Rhodes, 1980), (Wang et al., 2020), (Hazineh et al., 2021). Here we present a particularly simple protocol to perform this task, based on a camera where low-pass-filtering directional photodetectors are

combined with standard pixels in a checkerboard pattern. The same approach can be extended to implement more complex incoherent filtering operations that similarly require suppression of the low-frequency components, e.g., for object recognition. Altogether, we believe that these results will motivate and guide the development of a new class of image sensors based on metasurface technology, specifically designed to address a wide range of advanced image processing tasks with highly miniaturized form factor and reduced computational load.

2.3 Design and Simulations

The physical structure and principle of operation of the devices employed in this work are illustrated in Figure 2.1(a). In these devices, the illumination window of a photodetector is coated with a composite metasurface consisting of an optically thick metal film (Au) stacked with a periodic array of rectangular Au nanoparticles (NPs) and perforated with sub-wavelength slits. Two dielectric layers (SiO_2) are also introduced immediately below and above the metal film, to provide electrical insulation from the active layer and to control the film-NP coupling, respectively. The two SiO_2 layers have a nominal thickness of 60 nm, the metal film consists of 5 nm of Ti and 100 nm of Au, and each NP consists of 5 nm of Ti and 50 nm of Au with a width of 250 nm. These parameters of the grating platform are obtained by FDTD optimization, to be discussed in the next paragraph. Light incident at the desired detection angle is diffracted by the NPs into SPPs at the top surface of the metal film. These guided waves are then scattered by the slits into radiation propagating predominantly into the absorbing active layer, similar to the phenomenon of extraordinary optical transmission through sub-wavelength apertures in metal films (Ebbesen et al., 1998). Each slit section contains 5 slits with 200-nm width and 400-nm center-to-center spacing. Light incident along any other direction is instead either reflected or diffracted back into the air above. Specifically, in the present work we use devices where the array is surrounded symmetrically by slits on both sides, leading to a symmetric angular response peaked either at normal incidence (device A) or at equal and opposite illumination angles (device B) depending on the array

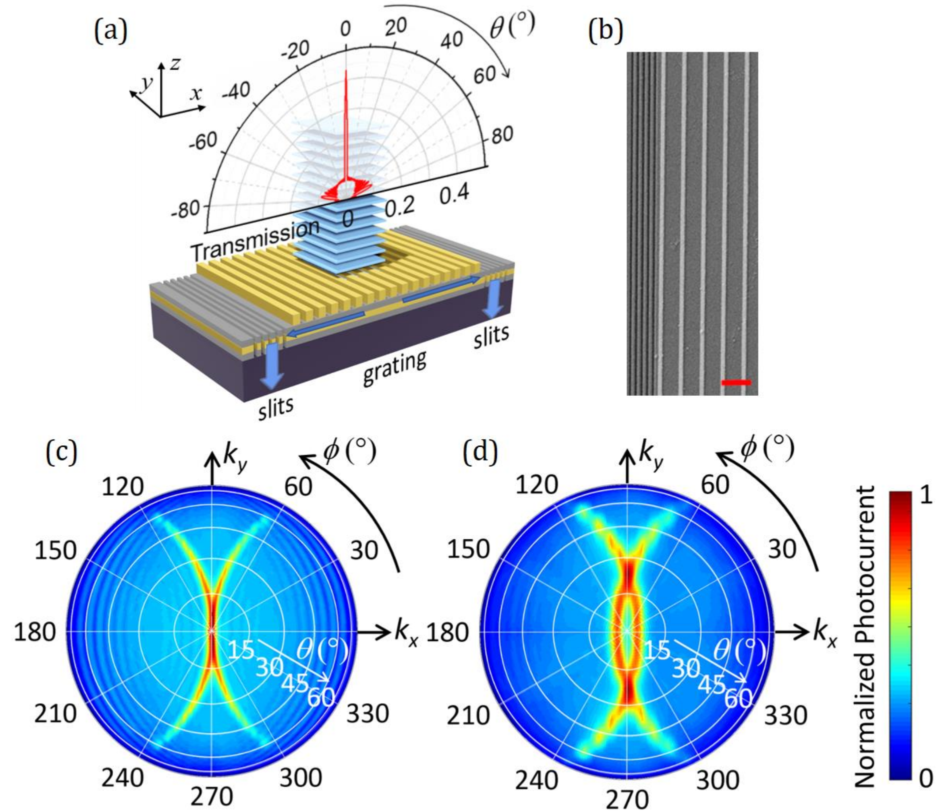

Figure 2.1: Plasmonic directional image sensors. (a) Schematic illustration of the physical structure and principle of operation for a device designed to provide angle-sensitive photodetection peaked at normal incidence. The polar plot shows the calculated optical transmission coefficient through the device metasurface for *p*-polarized light at λ_0 = 1550 nm versus angle of incidence θ on the x–z plane. (b) SEM image of an experimental sample (device A) showing a few periods of the NP array and the adjacent slits. The scale bar is 2 μm. (c) Measured responsivity of the same device versus polar θ and azimuthal ϕ illumination angles at λ_0 = 1550 nm, summed over two orthogonal polarizations. In this device, periodic array contains 15 NPs of 250 nm width with a period Λ = 1485 nm. (d) Same as (c) for a different sample (device B) featuring a symmetric double-peaked angular response with maximum photocurrent at θ = ±3.8°. The geometrical parameters of this device are nominally the same as in the sample of (c), except for a larger array period Λ = 1581 nm.

period Λ. In passing, we note that the same design platform can also be used to produce an asymmetric angular response peaked at any desired off-axis angle, by replacing the slits on one side of the array with a suitable "reflector" unit (Kogos et

al., 2020). While this configuration is not considered in the present work, it opens up additional spatial-filtering opportunities in the context of phase-contrast imaging, where asymmetric transfer functions are particularly beneficial (Tian and Waller, 2015). This work will be presented in the next chapter.

The polar plot of Figure 2.1 (a) shows the *p*-polarized transmission coefficient at 1550-nm wavelength of the metasurface of device A, computed as a function of polar illumination angle θ on the x-z plane with FDTD simulations. The Ansys Lumerical FDTD software package is used throughout this Dissertation. To obtain this line plot, I conducted two-dimensional simulations with Bloch boundary conditions on the lateral boundaries enclosing a full pixel that comprises one grating section plus a set of slits (each pixel shares a group of slits with the next adjacent one). In these calculations, the pixel is illuminated with a p-polarized plane wave and the transmitted light intensity into the device substrate is calculated for different values of the angle of incidence θ on the x-z plane. The device geometrical parameters are listed in the figure caption. A sharp transmission peak centered at $\theta = 0°$ (normal incidence) is observed in this plot, with full width at half maximum as small as 3° and peak value above 45%, originating from the excitation of SPPs propagating towards both sets of slits. For *s*-polarized incident light, the calculated transmission coefficient through the same metasurface is isotropic and significantly smaller, < 0.2% at all angles, consistent with the polarization properties of SPPs (Kogos et al., 2020). The angular response of Figure 2.1(a), rescaled by a factor of about 0.5, therefore also applies to unpolarized illumination. Because of its reliance on

diffraction, the device operation is also intrinsically wavelength dependent, and monochromatic light at λ_0 = 1550 nm is considered throughout this work, which can be approximated even under broadband illumination with the addition of a spectral filter in front of the image sensor array. It should be noted, however, that both polarization-insensitive and achromatic broadband operation are also possible with the same general platform, by replacing the periodic NP array with a gradient metasurface (Ding et al., 2017) or multilevel diffractive elements (Banerji et al., 2019) and leveraging the enhanced design flexibility of such systems.

2.4 Device Fabrication and Measurement

Our experimental samples consist of metal-semiconductor-metal Ge photo-conductors, with the metasurfaces just described patterned in the region between two metal contacts deposited on the top surface of a Ge substrate. Such photoconductors are particularly simple to fabricate, and the same results in terms of angular response can be expected with any other type of photodetectors (including image-sensor photodiodes) regardless of their operation principle. The metasurfaces were developed with a multi-step fabrication process, including EBL for the slits and NPs [Figure 2.1(b)], as described below.

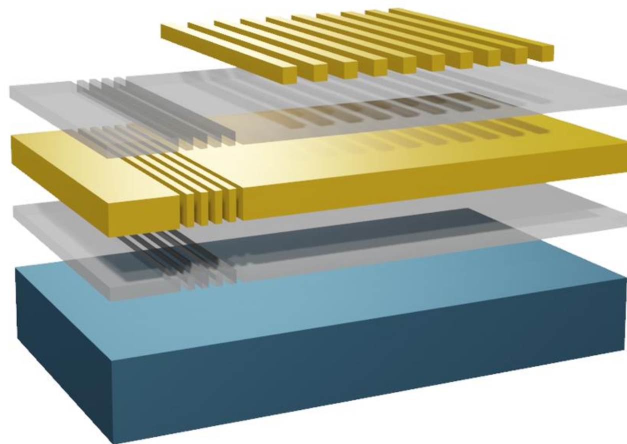

Figure 2.2: 3D schematic of device structure. The layers are deposited and patterned from the bottom to the top.

The experimental samples are fabricated on undoped (100) Ge substrates. The Au films (with a 5-nm Ti adhesion layer) are deposited by electron-beam evaporation, whereas RF sputtering is used for the SiO_2 layers. The Au film is patterned by photolithography to form the central device area and the two surrounding contacts. The SiO_2 films feature suitably positioned apertures to allow for device biasing and

current extraction. The slits are defined by EBL and RIE with a positive/negative double layer of PMMA and HSQ resist (Kogos et al., 2020) followed by deposition of the Ti/Au film and liftoff. The Au nanostripes are patterned by EBL with a single positive resist (PMMA). Therefore, at the device area we obtain a dielectric-metal-dielectric-metal sandwich, as shown by Figure 2.2. The experimental samples consist of a few (7) identical repetitions of the structure of Figure 2.1(a) where each section shares the slits of the adjacent section, and with a large (300 μm) separation between the two electrodes. This arrangement (equivalent to multiple identical pixels binned together) is convenient for the angle-resolved device characterization, because it alleviates the need for tightly focused incident light that would degrade the measurement angular resolution. In the final step of the fabrication process, a Ti rectangle with an opening over the entire metasurface is deposited on the top SiO_2 layer and patterned by photolithography. This cover is introduced to suppress any spurious photocurrent that may otherwise be caused by light absorbed near the electrodes away from the metasurface. The completed device is then mounted on a copper block and wire-bonded to two Au-coated ceramic plates.

The completed devices were then characterized by measuring their photocurrent under laser light illumination at 1550-nm wavelength as a function of polar θ and azimuthal ϕ angles of incidence, with a custom-built optical goniometer setup (Figure 2.3). The device under study is biased with a 1-V dc voltage and illuminated with 0.5-mW linearly polarized light from a diode laser. The incident optical power is modulated at 1 kHz, so that the photocurrent can be measured separately from the

dark current at low noise using a bias tee and lock-in amplifier. The laser light is delivered to the device with a polarization-maintaining fiber mounted in a cage system, which is rotated with a piezo-controlled stage about the focal point of its output lens to vary the polar angle of incidence θ. The device is also mounted on another rotational stage that allows tuning the azimuthal illumination angle ϕ. The polar angle is varied between ±85° in steps of 1°, whereas the measured azimuthal angles range from 0° to 90° in steps of 5°. The remaining two quadrants of the angular response maps are filled in based on the mirror symmetry of the device geometry. Finally, a linear interpolation is used to include additional data points between the measured values of ϕ in steps of 1° (MATLAB code see Appendix A).

The detailed system setup and measurement protocol were developed by Leonard Kogos (as was introduced in his Dissertation, see Figure 2.3), and I updated the alignment procedure for improved convenience and higher precision.

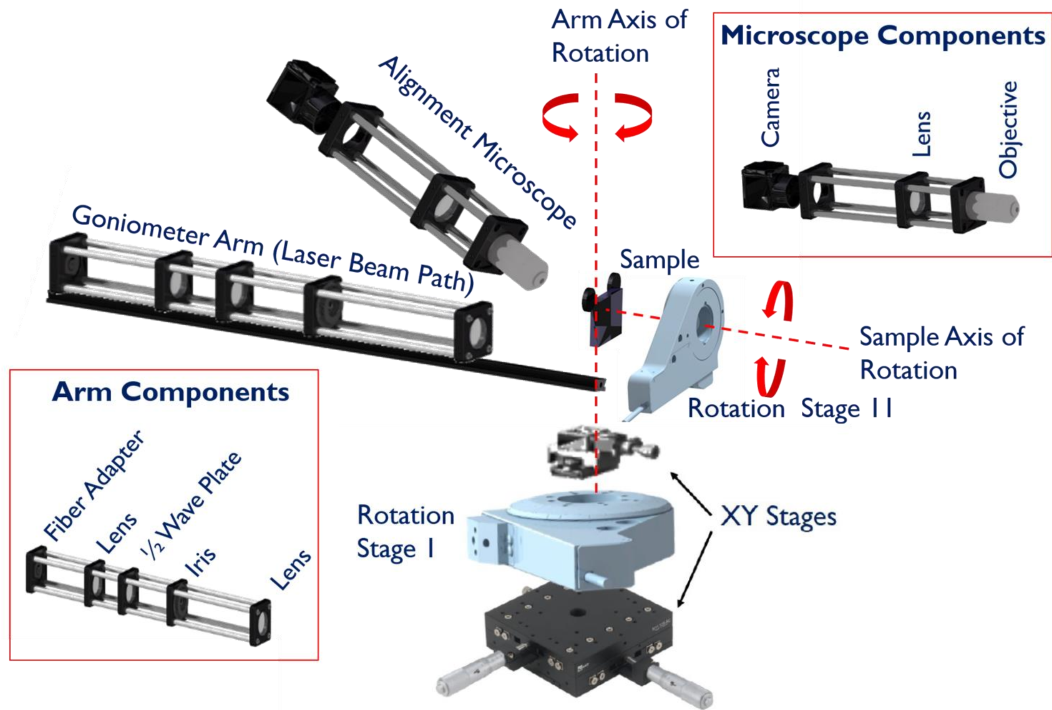

Figure 2.3: Illustration of the device measurement setup with major optical components. The insets show details about the composition of the goniometer arm and custom alignment microscope.

The experimental results [Figure 2.1(c) and (d)] show highly directional response in good agreement with theoretical expectations. In particular, the incident directions of high responsivity form a rather narrow distribution within the full hemisphere, consisting of two C-shaped regions of opposite curvature. The shape of this distribution is determined by the diffractive coupling of the incident light into different SPP modes, and the two C-shaped regions correspond to SPPs collected by the two slit sections surrounding the NP array. In sample A [Figure 2.1(c)], these two regions overlap at $\theta = 0°$ so that a single peak is produced in the horizontal line cut of the angular response. This device can therefore provide low-pass spatial filtering along the x direction. In contrast, in sample B the two C-shaped regions are slightly

offset from one another around $\theta = 0°$ [Figure 2.1(d)], leading to two symmetrically located response peaks at $\theta = \pm 3.8°$. In conjunction with an imaging lens of suitably small numerical aperture, the resulting transfer function corresponds to a high-pass filter. The experimental responsivities of these and similar devices were also compared to reference samples without any metasurface. The results are generally consistent with the calculated metasurface transmission penalty (about 45% and 23% for p-polarized and unpolarized light, respectively, as discussed above), although large sample-to-sample variations were observed (even in the reference samples) due to fabrication imperfections.

2.5 Coherent Transfer Function and Phase Contrast Imaging

In order to establish a connection between the angular response maps of Figure 2.1(c) and (d) and the spatial filtering capabilities of the same devices, here we introduce and evaluate the corresponding filter transfer function. In a standard optical spatial filter, such as a 4f system or nanophotonic equivalent, the input and output signals are the optical field distributions $E_{in}(\mathbf{r})$ and $E_{out}(\mathbf{r})$ on the input and output planes, respectively, and the CTF is defined as the ratio of their Fourier transforms $t(\mathbf{k}) = E_{out}(\mathbf{k})/E_{in}(\mathbf{k})$. In contrast, an array of directional plasmonic image sensors converts its incident optical field distribution $E_{in}(\mathbf{r})$ into a plasmonic field distribution $E_{SPP}(\mathbf{r})$, which is then sampled at the slit locations through the slit-scattering/photodetection process illustrated in Figure 2.4.

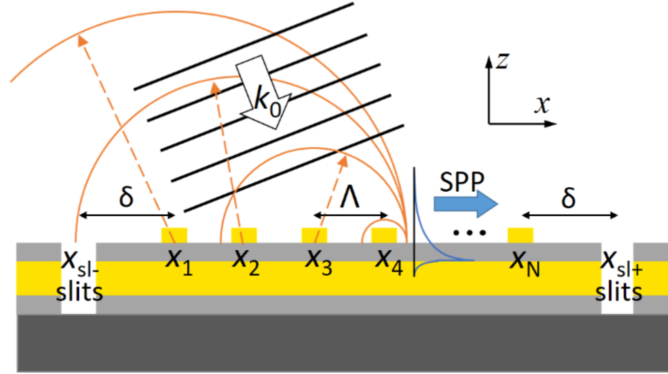

Figure 2.4: Schematic illustration of the physical model used to evaluate the CTF contribution $t_+(\mathbf{k})$ from the slits on the right of the NP array. The circles illustrate the phase relationship among the light waves scattered by different NPs into SPPs.

With this in mind, we can take $E_{SPP}(\mathbf{r})$ as the output signal of interest $E_{out}(\mathbf{r})$, with the understanding that such signal is only meaningfully defined at the slit locations $\mathbf{r} = \mathbf{r}_{sl}^{\boldsymbol{n}}$, because it cannot be measured by the sensor array anywhere else (here $\boldsymbol{n} =$

$\{n_x, n_y\}$ denotes a pair of integers n_x and n_y that label the different pixels in the array, and the spatial variable $\mathbf{r}_{sl}^{\boldsymbol{n}}$ indicates position along the slits in the $\boldsymbol{n}^{th}$ pixel). A CTF can then again be defined as the ratio of the Fourier transforms of the (discrete-space) output and input signals $E_{out}(\mathbf{r}_{sl}^{\boldsymbol{n}}) = E_{SPP}(\mathbf{r}_{sl}^{\boldsymbol{n}})$ and $E_{in}(\mathbf{r}_{sl}^{\boldsymbol{n}})$, and finally used to compute the image recorded by the sensor array.

To evaluate this CTF, we begin by considering the model structure shown in Figure 2(a), which contains N NPs arranged periodically at positions x_l (l = 1, 2, 3, …, N) with period $\Lambda = x_l - x_{l-1}$, and one slit located to the right of the NPs at x_{sl+} (for simplicity in this discussion we omit the pixel-label superscripts $\boldsymbol{n}$). The device is illuminated with a harmonic plane wave of in-plane wavevector component $k = (2\pi/\lambda_0)\sin\theta$ along the x direction, so that the incident field distribution on the NP array is $E_{in,k}(x) = E_{in}(k)e^{ikx}$. The incident light is scattered by all NPs, and the scattered waves can excite SPPs on the underlying metal film if the requirements of energy and momentum conservation are satisfied. The resulting SPP field at the slit position can then be expressed as

$$E_{SPP,k}(x_{sl+}) = \sum_{l=1}^{N} E_{in,k}(x_l) \int d\tilde{k}\, \eta_{SPP+}(k + \tilde{k}) e^{i(k+\tilde{k}+i\gamma)(x_{sl+}-x_l)}, \qquad (2.1)$$

where each term in the sum is the contribution from a different NP. In the integral, $\eta_{SPP+}(k + \tilde{k})$ is the probability amplitude that the light scattered by each NP with in-plane wavevector $k + \tilde{k}$ excites a SPP that propagates in the +x direction (i.e., towards the slit at x_{sl+}). The exponential factor accounts for the phase shift and attenuation experienced by this SPP as it travels from the NP to the slit. The SPP propagation

losses due to absorption and scattering are modeled with an attenuation coefficient $\gamma = 1/(2L_{SPP})$, where L_{SPP} is the SPP propagation length. The k dependence of η_{SPP+} is determined by the phase matching condition including the SPP lifetime broadening, so that $|\eta_{SPP+}|^2$ can be expressed as a Lorentzian function of $k + \tilde{k}$ centered at k_{SPP} (the SPP wavenumber at the illumination wavelength λ_0) with full width at half maximum (FWHM) $1/L_{SPP}$.

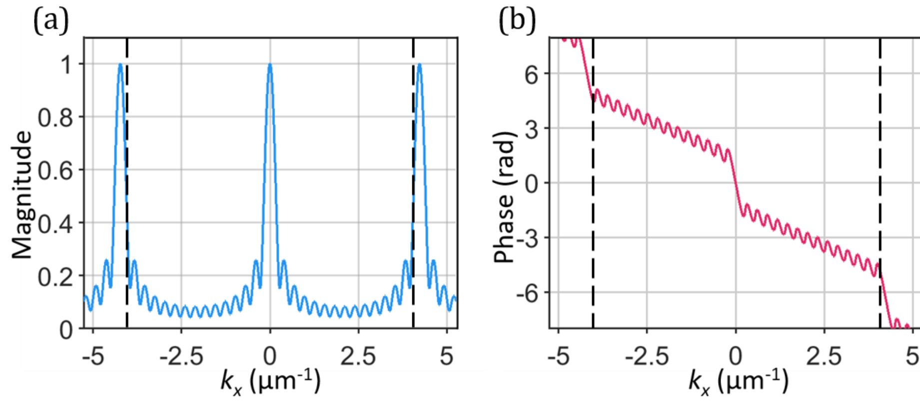

Figure 2.5: Coherent transfer function of the plasmonic directional image sensors of Figure 2.1. (a), (b) Magnitude (normalized to unit peak value) (a) and phase (b) of $t_+(k_x,k_y=0)$ versus k_x, computed using Equations (2.3) and (2.4) with the parameter values of sample A. The dashed vertical lines indicate the range of k values (from $-2\pi/\lambda_0$ to $+2\pi/\lambda_0$) accessible with external illumination from air. The corresponding plots for the CTF contribution $t_-(\mathbf{k})$ from the slits on the left of the NP array can be inferred from these traces using the relation $t_-(\mathbf{k}) = t_+(-\mathbf{k})$.

Equation (2.1) can be simplified by using $E_{in,k}(x_l) = E_{in}(k)e^{ikx_l}$ and $x_{sl+} - x_l = (N - l)\Lambda + \delta$, where $\delta = x_{sl+} - x_N$ is the distance between the slit and its nearest NP [see Figure 2.1(a)]. With these substitutions, we find that the SPP field at the slit can be expressed as $E_{SPP,k}(x_{sl+}) = t_+(k)E_{in,k}(x_{sl+})$, where

$$t_+(k) = \int d\tilde{k}\, \eta_{SPP+}(k + \tilde{k})e^{i(\tilde{k}+i\gamma)\delta}f(\tilde{k}) \tag{2.2}$$

and

$$f(\tilde{k}) = \sum_{l=1}^{N} e^{i(\tilde{k}+i\gamma)(N-l)\Lambda} = \left[1 - e^{i(\tilde{k}+i\gamma)N\Lambda}\right]/\left[1 - e^{i(\tilde{k}+i\gamma)\Lambda}\right]. \qquad (2.3)$$

According to these equations, the output signal $E_{SPP,k}(x_{sl+})$ sampled by the plasmonic image sensor under plane-wave illumination is linearly related to the input field $E_{in,k}(x_{sl+})$ at the same location, as in a traditional optical spatial filter. The k-dependent proportionality factor $t_+(k)$ is therefore the contribution to the device CTF from the slit at x_{sl+}. The same analysis can be readily extended to evaluate the contribution $t_-(k)$ from a slit located symmetrically on the left-hand side of the NPs (at position $x_{sl-} = x_1 - \delta$), and to include a finite y component for the in-plane wavevector **k** of the incident light. The resulting expression for the CTF is

$$t_\pm(\mathbf{k}) = \int d\tilde{k}\, \eta_{SPP\pm}(\mathbf{k} \pm \hat{\mathbf{x}}\tilde{k}) e^{i(\tilde{k}+i\gamma)\delta} f(\tilde{k}), \qquad (2.4)$$

where η_{SPP+} and η_{SPP-} describe the excitation of SPPs propagating in the positive and negative x directions, respectively, and therefore account for the two C-shaped regions of high responsivity observed in the angular response maps of these devices [see Figure 2.1(c) and (d)]. Multiple pairs of symmetrically positioned slits [as in the structure of Figure 2.1(a)] can be modeled in the same fashion, resulting in the same expression for $t_\pm(\mathbf{k})$ with slightly different values of δ. However, as long as the separation between adjacent slits is small compared to the pixel size, inclusion of these different values has negligible effect on the overall frequency response.

The CTF $t_\pm(k_x,k_y=0)$ of Equations (2.3) and (2.4) consists of a series of identical peaks centered at $k_x = k_{SPP} - mg$, where $g = 2\pi/\Lambda$, m is an integer, and each peak

corresponds to a different order of diffraction. Figure 2.5(a) and (b) show, respectively, the magnitude (normalized to unit peak value) and phase of $t_+(k_x,k_y=0)$ computed with these equations. The corresponding plots for $t_-(\mathbf{k})$ can be inferred directly from these traces using the relation $t_-(\mathbf{k}) = t_+(-\mathbf{k})$, which follows from Equation (2.4) and is consistent with the symmetric device geometry under study. These calculations are based on the parameter values of sample A, including $\Lambda = \delta =$ 1485 nm, N = 15, λ_0 = 1550 nm, and $k_{SPP} = 2\pi/\Lambda$ so that the *m*=1 peak is centered at $k_x = 0$ as in Figure 2.1(c). For the SPP propagation length we use L_{SPP} = 80 μm, selected with a numerical fit so that the peaks of $|t_\pm(k_x,k_y=0)|^2$ have the same linewidth as in our measured responsivity data [the horizontal line cut of the color map in Figure 1(c)]. The dashed vertical lines in Figure 2.5 indicate the range of k values (from $-2\pi/\lambda_0$ to $+2\pi/\lambda_0$) accessible with external illumination from the air above the NP array. The experimental data of Figure 2.1(c) are well reproduced by the calculation results plotted in Figure 2.5(a), including the fringes around the main peak which originate from incomplete cancellation of the scattered waves away from the Bragg condition in the presence of a finite number of NPs.

As shown in Figure 2.5(b), the phase response $\phi_+(\mathbf{k}) = \arg\{t_+(\mathbf{k})\}$ exhibits a linear dependence on k_x with negative slope $d\phi_+(\mathbf{k})/dk_x = -\alpha$ across the entire linewidth of the peak at $\mathbf{k} = 0$, i.e., for all accessible values of $\mathbf{k}$ for which $|t_+(\mathbf{k})|$ is non-negligible. The value of the slope parameter α inferred from this plot is 9.8 μm, which is relatively close to the distance between the slit and the center of the NP array $x_{sl+} - x_c$

= 11.9 μm in device A. In fact, α becomes exactly equal to $x_{sl+} - x_c$ in the limit of large L_{SPP}. Additional FDTD simulations similarly indicate a linear phase profile for $t_+(k_x,k_y=0)$ near $k_x = 0$ with comparable slope parameter α = 12.2 μm [the smaller value of 9.8 μm in Figure 2.5(b), where L_{SPP} is obtained from a fit to the experimental data, suggests a shorter SPP propagation length in the actual samples compared to the FDTD simulations, likely due to SPP scattering by surface roughness]. By the shifting property of Fourier transforms, this linear phase profile corresponds to a displacement in real space by the amount α in the negative x direction. We can therefore conclude that the SPP signal sampled by the slit at x_{sl+} is an amplitude-filtered version of the light incident on the device at position $x_{sl+} - \alpha$, close to the center of the NP array. Similar considerations apply to the phase of $t_-(\mathbf{k})$ with positive slope α, so that the slit at x_{sl-} filters the input signal at $x_{sl-} + \alpha$, also close to the center of the device. It should be noted that this sampling behavior is fundamentally different from the operation of standard photodetectors, which instead average the incident light across their entire illumination window.

To evaluate the filtering capabilities of our experimental samples, in the calculations presented below the CTF is expressed as

$$t_{\pm}(\mathbf{k}) \propto e^{\mp i\alpha k_x}\sqrt{R_{\pm}(\mathbf{k})}, \tag{2.5}$$

where $R_{\pm}(\mathbf{k})$ is the contribution to the measured angle-resolved responsivity from the slits at $x = x_{sl\pm}$. This formula follows from the observation that $R_{\pm}(\mathbf{k})$ is proportional to the magnitude squared of the SPP field at the slit locations, with a **k**-

independent proportionality factor determined by the efficiency of the SPP slit-scattering process and the quantum efficiency of the photodetector active layer. For the phase slope parameter α, we use the value of 9.8 μm obtained from the analytical model above fitted to the experimental data. When the same devices are illuminated with an arbitrary incident field $E_{in}(\mathbf{r}) = \int d\mathbf{k}\, E_{in}(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{r}}$, the measured photocurrent is proportional to the sum of the intensities of the corresponding output signals measured by the two slits $E_{SPP}(\mathbf{r}_{sl\pm}) = \int d\mathbf{k}\, t_{\pm}(\mathbf{k})E_{in}(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{r}_{sl\pm}}$, each averaged over the slit length along the y direction. In these calculations, the CTF $t_{\pm}(\mathbf{k})$ can again be evaluated using Equation (2.5). Furthermore, the Fourier transform of the incident light $E_{in}(\mathbf{k})$ can be related to that of the object $E_{obj}(\mathbf{k})$ according to $E_{in}(\mathbf{k}) = t_{lens}(\mathbf{k})E_{obj}(\mathbf{k})$, where $t_{lens}(\mathbf{k})$ is the transfer function of the imaging lens in front of the sensor array. Under coherent illumination, $t_{lens}(\mathbf{k})$ is a cylindrical step function with cutoff frequency $k_c = \pi/(\lambda_0 F)$, where F is the lens F number. With these prescriptions, we can compute the image of any object produced by any array of plasmonic directional sensors.

As an example, let's consider a transparent phase object [the phase-grating set shown in Figure 2.6(a)] imaged with a camera where every pixel is coated with the metasurface of device B. In this case, the object field $E_{obj}(\mathbf{r})$ has uniform amplitude across the entire field of view, and therefore the object could not be visualized using a standard imaging system with **k**-independent response, except for negligibly small diffraction fringes [Figure 2.6(b)]. At the same time, as light is transmitted through the object, its local direction of propagation is deflected by an angle proportional to

the local phase gradient. As a result, if the system response varies with angle of incidence, the recorded signal distribution acquires a dependence on the object phase gradient, and therefore spatially varying features of the object (in this case, the edges of the phase-grating lines) can be resolved. The same behavior can also be described in the spatial frequency domain as edge enhancement caused by the CTF **k** dependence. These ideas have been explored extensively with different types of optical spatial filters (Zernike, 1942), (Kim et al., 2012), (Fürhapter et al., 2005), (Kwon et al., 2020), for applications ranging from label-free imaging of biological samples (Park et al., 2018) to semiconductor wafer inspection (Zhou et al., 2013). By virtue of their intrinsic angular sensitivity, the directional image sensors under study can provide the same functionality without any external filtering elements.

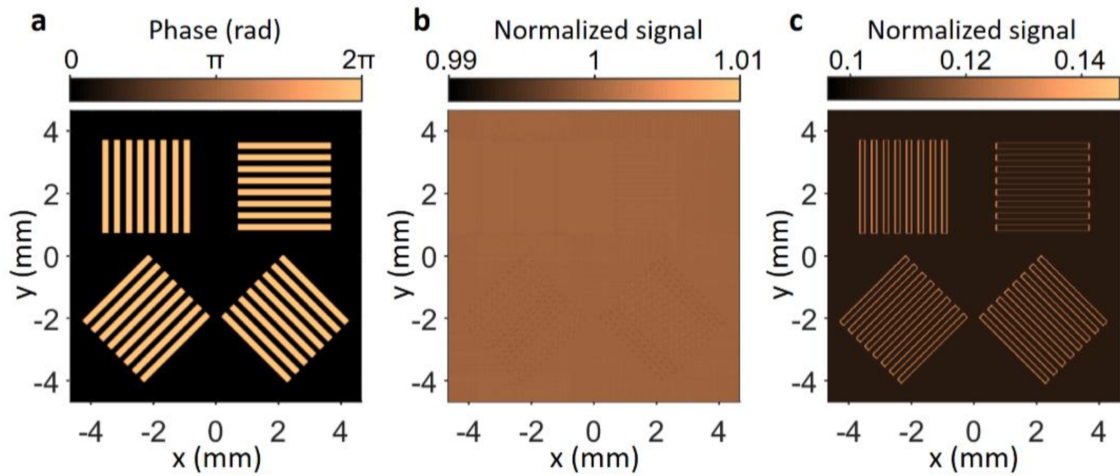

Figure 2.6: Phase imaging results. (a) Phase distribution of the object. (b) Image of the object of (a) computed for an array of 392×392 uncoated pixels [i.e., with **k**-independent CTF] combined with an NA=0.13 imaging lens. (c) Image of the same object computed for an otherwise identical camera where every pixel is coated with the metasurface of device B [modeled using the data of Figure 2.1(d)]. The signal intensity in (b) and (c) is normalized to that of the uncoated devices when illuminated with the same plane wave incident on the object.

To illustrate, I have computed the image of the phase gratings of Figure 2.6(a) for an array of 392×392 pixels described by the experimental angular response map of device B [Figure 2.1(d)], combined with an F/3.8 imaging lens (corresponding to a numerical aperture NA = 0.13 and a field of view of 15°). In particular, following the prescriptions above, this image was obtained by summing the contributions from the two slit sections governed by the CTFs $t_{\pm}(\mathbf{k})$. The resulting plot is shown in Figure 2.6(c), where the signal measured by each plasmonic pixel is normalized to that of an identical uncoated device under the same illumination conditions. The y-oriented edges of the grating lines are clearly visualized in this image, consistent with the strong k_x-dependence of the responsivity $R(\mathbf{k})$ of device B at small angles of incidence. The x-oriented edges can also be discerned, but with significantly lower contrast, due to the weaker variations of $R(\mathbf{k})$ with k_y (mostly related to the C shapes of the responsivity peaks). For comparison, the grating lines are essentially invisible in the image computed for an otherwise identical camera of standard pixels with $\mathbf{k}$-independent CTF [Figure 2.6(b)]. It should also be noted that isotropic phase imaging could similarly be achieved with this approach, using alternative metasurface designs featuring rotationally-invariant angular response, e.g., based on circular NPs and slits.

2.6 Optical Transfer Function and Optical Spatial Filtering

Next, we consider the frequency response of the same devices under natural, and therefore spatially incoherent, illumination. In this case, the incident light features a highly localized correlation function, which can be modeled as $\langle E_{in}^*(\mathbf{r}-\frac{\delta r}{2})E_{in}(\mathbf{r}+\frac{\delta r}{2})\rangle \propto I_{in}(\mathbf{r})\exp(-\frac{\delta r^2}{4\Delta^2})/\Delta$, where the brackets $\langle\ldots\rangle$ indicate an ensemble average and the transverse coherence length Δ is small compared to the size of the image. Under these conditions, the operation of any optical spatial filter is governed by its OTF: $T(\mathbf{q}) = I_{out}(\mathbf{q})/I_{in}(\mathbf{q})$, where $I_{in}(\mathbf{q})$ and $I_{out}(\mathbf{q})$ are the Fourier transforms of the input and output field-intensity distributions, respectively (Goodman, 2005). This function can be computed by expressing the output intensity $I_{out}(\mathbf{r}) \propto \langle E_{out}^*(\mathbf{r})E_{out}(\mathbf{r})\rangle$ in terms of the Fourier transform of the output field $E_{out}(\mathbf{k}) = t(\mathbf{k})E_{in}(\mathbf{k})$, and then evaluating the ensemble average in the spatial-frequency domain. In the directional image-sensor arrays of interest, $I_{out}(\mathbf{r})$ is again only accessible at the slit locations $\mathbf{r}_{sl\pm}^{\boldsymbol{n}}$, but otherwise the OTF can be defined and computed in the same fashion. After a derivation with the application of the Gaussian model $\langle E_{in}^*(\mathbf{r}-\frac{\delta r}{2})E_{in}(\mathbf{r}+\frac{\delta r}{2})\rangle \propto I_{in}(\mathbf{r})\exp(-\delta r^2/4\Delta^2)/\Delta$, the resulting expression is

$$T_{\pm}(\mathbf{q}) \propto \int d\mathbf{k}\, t_{\pm}^*(\mathbf{k})t_{\pm}(\mathbf{k}+\mathbf{q})\exp\left(-\frac{\Delta^2|2\mathbf{k}+\mathbf{q}|^2}{4}\right), \quad (2.6)$$

with $T_-(\mathbf{q}) = T_+^*(\mathbf{q}) = T_+(-\mathbf{q})$ in a symmetric device. In the limit where $\Delta \to 0$ (i.e., for completely incoherent light), $T_{\pm}(\mathbf{q})$ is simply equal to the autocorrelation function of the CTF $t_{\pm}(\mathbf{k})$, which is maximum for $\mathbf{q} = 0$ regardless of the detailed wavevector dependence of $t_{\pm}(\mathbf{k})$. This observation confirms the aforementioned statement that,

under incoherent illumination, an optical spatial filter cannot be used to perform any operation that requires suppression of the DC components of the image, such as edge detection by spatial differentiation.

Figures 2.7(a) and (b) show the magnitude of the OTF $T_+(\mathbf{q})$ of devices A and B, respectively, computed with the model just presented. In these calculations, the parameter Δ of Equation (2.6) is evaluated based on the Van Cittert – Zernike theorem (Goodman, 2005). Specifically, we use Δ = 6.7 μm, which corresponds to a fully incoherent object at a representative distance of 10 times its lateral size. It should be noted that the corresponding radius of coherence (18.9 μm) is sufficiently large to ensure spatial coherence of the incident light across the entire NP array of each device. Furthermore, this value of Δ should be regarded as a lower bound, because it neglects the finite coherence of the light at the object. The CTF $t_\pm(\mathbf{k})$ in Equation (2.6) is computed using Equation (2.5), with the responsivity contributions from the two slits $R_\pm(\mathbf{k})$ obtained from the experimental data of Figures 2.1(c) and (d). As expected, both transfer functions plotted in Figure 2.7 are nonzero and near-maximum at $\mathbf{q}$ = 0, which again originates from the autocorrelation nature of Equation (6). At the same time, the detailed shape of these OTFs is determined by the **k**-dependence of the corresponding CTFs (which in turn can be tailored through the metasurface design), combined with the windowing action of the exponential term in Equation (6). More complex OTFs can therefore be envisioned, for example involving

additional peaks at finite **q** values, with metasurfaces designed to produce multiple peaks in $t_+(\mathbf{k})$ and $t_-(\mathbf{k})$ individually.

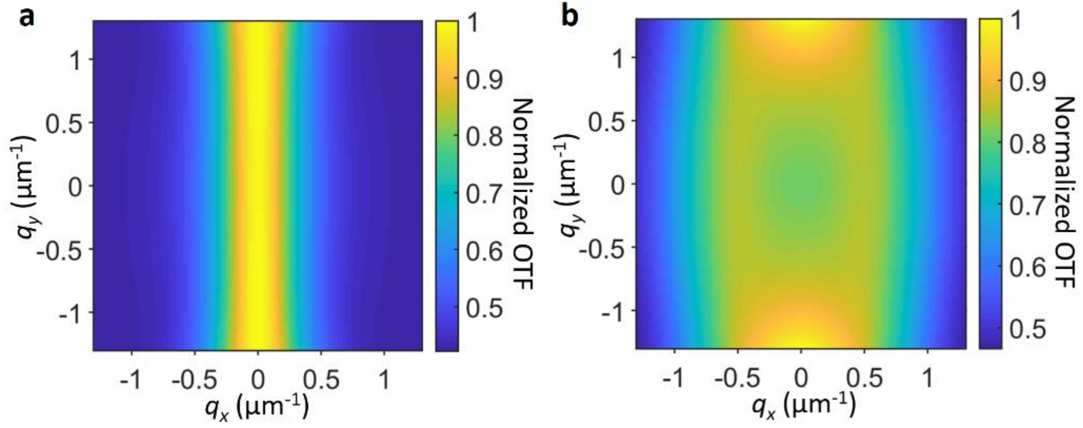

Figure 2.7: Optical transfer function of the plasmonic directional image sensors of Figure 2.1. Panels (a) and (b) show the magnitude of $T_+(\mathbf{q})$ for devices A and B, respectively, computed using their measured responsivity maps and normalized to unit peak value.

The transfer function of Equation (2.6) can be used to model the incoherent imaging capabilities of arrays of these plasmonic devices. To that purpose, we begin by noting that the photocurrent measured by the $\boldsymbol{n}^{\text{th}}$ pixel is proportional to $\mathrm{I}_{\text{meas}}(\boldsymbol{n}) = \mathrm{I}_{\text{meas}+}(\boldsymbol{n}) + \mathrm{I}_{\text{meas}-}(\boldsymbol{n})$, where

$$\mathrm{I}_{\text{meas}\pm}(\boldsymbol{n}) = \int_{-\mathrm{w}/2}^{\mathrm{w}/2} \frac{\mathrm{dy}}{\mathrm{w}} \mathrm{I}_{\text{out}\pm}(\mathrm{x}_\mathrm{c}^{\boldsymbol{n}} \pm \mathrm{d}, \mathrm{y}_\mathrm{c}^{\boldsymbol{n}} + \mathrm{y}) \tag{2.7}$$

is the intensity detected by each slit, averaged over the slit length w along the y direction. Here, as before, $\boldsymbol{n} = \{n_\mathrm{x}, n_\mathrm{y}\}$ indicates a pair of integers n_x and n_y that label the different pixels, $\mathbf{r}_\mathrm{c}^{\boldsymbol{n}}$ is the center position of the $\boldsymbol{n}^{\text{th}}$ pixel, and $\mathrm{d} = \mathrm{x}_{\mathrm{sl}+}^{\boldsymbol{n}} - \mathrm{x}_\mathrm{c}^{\boldsymbol{n}} = \mathrm{x}_\mathrm{c}^{\boldsymbol{n}} - \mathrm{x}_{\mathrm{sl}-}^{\boldsymbol{n}}$ is the slit-to-center distance. Next, we express $\mathrm{I}_{\text{out}\pm}(\mathbf{r})$ in Equation (2.7) in terms of its Fourier transform $\mathrm{I}_{\text{out}\pm}(\mathbf{q}) = \mathrm{T}_\pm(\mathbf{q})\mathrm{I}_{\text{in}}(\mathbf{q})$, evaluate the integral over y, and

finally extract the Fourier transform of $I_{meas}(\boldsymbol{n}) = \int d\mathbf{q} I_{meas}(\mathbf{q}) e^{i\mathbf{q}\cdot\mathbf{r}_c^{\boldsymbol{n}}}$. With this procedure [and using $T_-(\mathbf{q}) = T_+^*(\mathbf{q})$] we find that $I_{meas}(\mathbf{q}) = T_{meas}(\mathbf{q}) I_{in}(\mathbf{q})$, where

$$T_{meas}(\mathbf{q}) = \mathrm{sinc}\left(\frac{q_y w}{2\pi}\right) \mathrm{Re}\left\{e^{i q_x d} T_+(\mathbf{q})\right\} \tag{2.8}$$

is an effective OTF that describes the measurement of the recorded image by the pixel array. In particular, pixelation effects are also included in this expression through its w and d dependence. Finally, the Fourier transform of the incident intensity $I_{in}(\mathbf{q})$ can be related to that of the object $I_{obj}(\mathbf{q})$ according to $I_{in}(\mathbf{q}) = T_{lens}(\mathbf{q}) I_{obj}(\mathbf{q})$, where $T_{lens}(\mathbf{q})$ is the OTF of the imaging lens:

$$T_{lens}(\mathbf{q}) = \begin{cases} \frac{2}{\pi}\left[\mathrm{acos}\left(\frac{q}{2k_c}\right) - \frac{q}{2k_c}\sqrt{1-\left(\frac{q}{2k_c}\right)^2}\right], & q < 2k_c \\ 0, & q > 2k_c \end{cases} \tag{2.9}$$

when we consider the CTF of the lens is a simple circle aperture. Under incoherent illumination $T_{lens}(\mathbf{q})$ decreases almost linearly from 1 to 0 as q varies from 0 to $2\pi/(\lambda_0 F)$.

The geometrical tunability of the transfer function of these devices is particularly significant for use in sensor arrays partitioned into identical blocks of multiple adjacent pixels, each coated with a different metasurface. With this arrangement, a single camera could produce multiple filtered images of a same object simultaneously, which could then be exploited to perform specific visual processing tasks (e.g., object recognition) with reduced electronic computational cost. Within this framework, even the constraint that $T_{meas}(\mathbf{q}=0) \neq 0$ under incoherent illumination can be effectively circumvented by subtracting the signals of different adjacent pixels

within each block, to produce an overall response equal to the difference of their respective OTFs. In fact, a similar idea has already been explored as a way to enable incoherent edge detection with a 4*f* system, where two masks of different cutoff frequencies are inserted sequentially at the Fourier plane and the resulting images are subtracted from one another (Rhodes, 1980). Such setup, however, is particularly bulky and rather impractical. More recently, nanophotonic implementations have also been proposed based on wavelength or polarization multiplexing (Wang et al., 2020), (Hazineh et al., 2021). By virtue of their ability to enable multiple filtering operations simultaneously on a pixel-by-pixel basis, the directional image sensors under study are ideally well suited to implement this general approach for incoherent image processing.

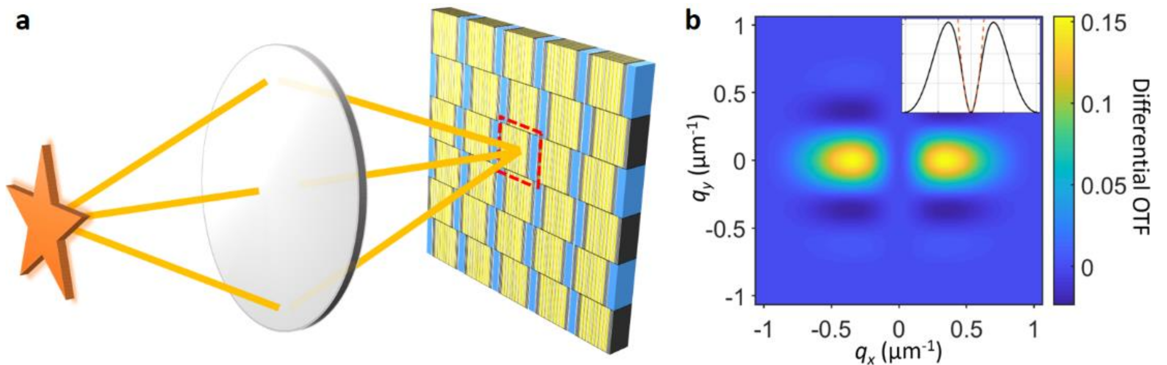

Figure 2.8: Incoherent edge detection protocol. (a) Schematic illustration of the envisioned sensor array, which consists of pixels based on the plasmonic directional photodetectors of Figure 2.1(a) (yellow squares) combined with bare pixels (blue rectangles) in a checkerboard pattern. The edge-enhanced image is obtained by subtracting the signals of neighboring pixels. (b) Differential OTF $\Delta T_{tot}(\mathbf{q})$ = $\Delta T_{meas}(\mathbf{q})T_{lens}(\mathbf{q})$ of the resulting imaging system with an NA=0.13 lens, computed by subtracting the cumulative OTFs of the two types of pixels. Here $\Delta T_{tot}(\mathbf{q})$ is normalized to the cumulative OTF of the bare pixels at $\mathbf{q}$ = 0. The inset shows the q_y = 0 line cut of $\Delta T_{tot}(\mathbf{q})$ (solid line), together with a numerical fit of the low-q_x portion of this trace to a quadratic function of q_x (dashed line).

As a particularly simple illustration, we consider the configuration shown schematically in Figure 2.8(a). Here, the low-pass-filter metasurface of device A is fabricated on every other pixel of an image sensor array in a checkerboard pattern, while all the other devices are left uncoated. The photocurrent measured by each plasmonic sensor is then subtracted from that of its adjacent uncoated pixel to produce a null at $\mathbf{q} = 0$, and therefore high-pass filtering. The resulting image is related to the object according to $I_{meas}(\mathbf{q}) = \Delta T_{tot}(\mathbf{q}) I_{obj}(\mathbf{q}) = \Delta T_{meas}(\mathbf{q}) T_{lens}(\mathbf{q}) I_{obj}(\mathbf{q})$, where $\Delta T_{meas}(\mathbf{q})$ is the difference between the measurement OTFs of the two neighboring pixels. In the uncoated reference pixels, the responsivity $R(\mathbf{k})$ is essentially constant with $\mathbf{k}$, and the frequency dependence of the measurement OTF is mostly determined by pixelation effects. At the same time, their photocurrent signal at normal incidence ($\mathbf{q} = 0$) is larger than in the plasmonic devices due to the aforementioned transmission penalty of the metasurfaces (about 23% for unpolarized light). This difference can be normalized out in the digital data processing before the subtraction step. Alternatively, it could be handled in the optical domain by reducing the size of the reference pixel along the x direction by the same factor of 23%, e.g., using the checkerboard pattern with alternating square and rectangular pixels shown in Figure 2.8(a). In fact, this approach provides several other important advantages. First, it reduces the size of each super-pixel (i.e., each block of adjacent coated and uncoated devices), which is favorable to increase the spatial resolution of the high-pass-filtered image. Second, it produces a flatter frequency response for the reference pixels across the full bandwidth of $T_{lens}(\mathbf{q})$, and

therefore increases the frequency range over which $\Delta T_{tot}(\mathbf{q})$ can be tailored through the plasmonic pixel design. Finally, it can also result in improved noise cancellation upon signal subtraction.

For the pixel configuration of Figure 2.8(a), this procedure leads to the differential OTF $\Delta T_{tot}(\mathbf{q})$ plotted in Figure 2.8(b). As expected, this transfer function is zero at $\mathbf{q} = 0$, and features two pronounced peaks at symmetric locations around the origin along the q_x direction. The imaging system of Figure 5(a) can therefore be used to enhance rapidly varying features of the object (i.e., edges) along the x direction. In particular, the $q_y = 0$ line cut of $\Delta T_{tot}(\mathbf{q})$ [shown by the solid line in the inset of Figure 2.8(b)] is well approximated by a quadratic function of q_x (dashed line) over a broad portion of the accessible spatial-frequency range. Since multiplication by q_x^2 in the frequency domain is equivalent to taking the second-order derivative with respect to x in the space domain, the pixel array of Figure 2.8 can provide (directional) spatial differentiation, which is a standard mathematical operation used for edge detection.

This behavior is illustrated in Figures 2.9(a) and (b), where we plot a simple amplitude object and the absolute value of its detected image, for an array of 171×210 super-pixels again combined with a lens of NA = 0.13. Clear edge enhancement is observed with maximum contrast for the edges oriented along the y direction, whereas x-oriented edges are not resolved. In Figure 2.9(c), the red and blue traces show, respectively, the line cut of the filtered image along the dashed line of Figure 2.9(b) and the corresponding object. The comparison between the two traces clearly demonstrates the second-order derivative nature of this optical spatial filter along

the x direction, which produces the two peaks per edge observed in the image. The magnitude of the detected signals in Figure 2.9(b) is limited by the differential nature of the underlying data acquisition, combined with pixelation effects. In any case, the resulting image can be fully resolved with the SNR levels accessible with near-infrared photo-detectors of similar dimensions (71 dB) (Murata et al., 2020). These results therefore demonstrate the feasibility of incoherent high-pass filtering with the plasmonic image sensors under study. Additional filtered versions of the same image (e.g., edge-enhanced along different orientations and/or with different transfer functions) could similarly be produced with the same general approach, using different metasurface designs and pixel arrangements.

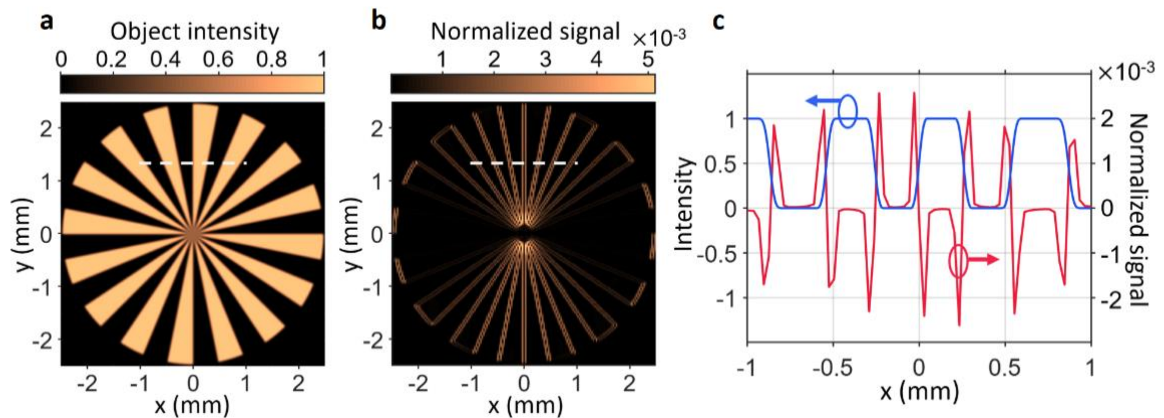

Figure 2.9: Incoherent edge detection results. (a), (b) Illustrative object (a) and absolute value of its edge-enhanced image (b) computed for an array of 171×210 super-pixels based on the configuration of Figure 2.8 combined with an NA=0.13 lens. (c) Red trace: line cut of the image along the dashed line of (b). Blue trace: line cut of the object along the same line. In (b) and (c), the differential signal produced by each super-pixel is normalized to the signal of the uncoated device under maximum illumination from the object.

2.7 Conclusion

In this chapter, I have introduced a new approach for optical spatial filtering based on pixel arrays of plasmonic directional image sensors with tailored angular response. To establish the image processing capabilities of these devices, we have developed a rigorous theoretical model of their filter transfer function under both coherent and incoherent illumination. The use of these devices for phase imaging and incoherent edge detection has also been demonstrated through imaging simulations based on the measured angle-resolved responsivity of prototype samples. These results are promising for a wide range of imaging applications including microscopy (e.g., for the visualization of transparent biological cells) and computer vision (e.g., for object recognition). Compared to more traditional optical spatial filters based on Fourier optics, this approach does not require any external filtering elements, and therefore can provide extreme size miniaturization and improved ease of alignment, which are beneficial for embedded and mobile applications. Furthermore, it allows controlling the filter transfer function on a pixel-by-pixel basis, so that multiple filtered images of a same object can be produced simultaneously, similar to the output of the first layer of a CNN. These images could then be fed into the subsequent CNN layers to perform various visual recognition tasks. Compared to fully electronic solutions, this ability to synthesize multiple filtered images in the optical domain can provide significant savings in power consumption, estimated at about tenfold based on prior studies of hybrid optoelectronic CNN configurations (Tian and Waller, 2015), (Chen et al., 2016).

CHAPTER THREE

Asymmetric Directional Photodetectors for Quantitative Phase Contrast Imaging

3.1 Introduction

3.1.1 Background

Traditional image sensors can only capture the intensity distribution of the incident light, whereas all information associated with the phase profile is lost in the image acquisition process. While these devices are clearly adequate for basic imaging tasks, direct access to the wavefronts and local directions of light propagation would allow for more advanced imaging capabilities. One example of particular interest is the ability to visualize phase-only objects where light is transmitted or reflected without any appreciable intensity variations. Relevant application areas where this capability plays a prominent role include microscopy for label-free imaging of biological samples (Park et al., 2018), surface profiling, and semiconductor inspection for detecting manufacturing defects (Zhou et al., 2013). Conventionally, phase imaging is achieved with rather complex and bulky setups, ranging from Zernike phase-contrast and differential-interference-contrast microscopy to quantitative techniques based on interferometry (Park et al., 2018) or non-interferometric methods (Paganin and Nugent, 1998), (Mehta and Sheppard, 2009). More recently, newly developed free-space nanophotonics and flat-optics platforms have also been applied to the demonstration of similar phase imaging systems, with the potential advantage of more compact dimensions and enhanced design flexibility (Pors et al., 2015), (Davis

et al., 2019), (Zhou et al., 2020), (Huo et al., 2020), (Kwon et al., 2020), (Zhou et al.,2021), (Ji et al., 2022), (Wang et al., 2023).

3.1.2 Principle of Design and Functionality

In this chapter, I will discuss a new type of image sensors that can measure the phase gradient of the incident optical field directly with the simplest possible setup, i.e., a standard camera or microscope without any external optical elements other than the imaging lenses. These devices consist of photodetectors individually coated with an integrated plasmonic metasurface that introduces a sharp dependence of responsivity $\mathcal{R}$ on illumination angle θ near normal incidence, on the same platform of design introduced in Chapter 2.

The wavefront sensing ability is illustrated schematically in Figure 3.1(a), where a plane wave of field amplitude $U_{in}(z) = U_0e^{ikz}$ is incident on a transparent object that introduces a position-dependent transmission phase shift $\varphi(x)$.

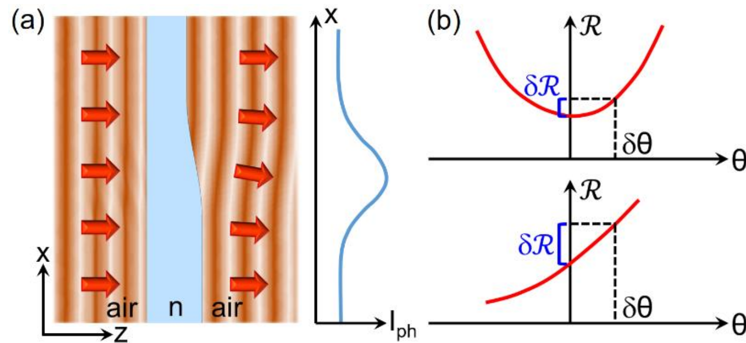

Figure 3.1: Phase contrast imaging with angle-sensitive photodetectors. (a) Left: wavefront distortion experienced by a plane wave after transmission through a transparent plate of variable thickness. Right: photocurrent signal I_{ph} measured by an angle-sensitive photodetector at different locations across the transmitted wavefront. (b) Responsivity $\mathcal{R}$ versus angle of incidence θ for a generic device with symmetric (top, Chapter 2) and asymmetric (bottom, this chapter) angular response. In the limit of small deflection $\delta\theta$, the asymmetric device provides a larger change in responsivity $\delta\mathcal{R}$, leading to increased image contrast.

Correspondingly, the direction of propagation of the transmitted wave $U_{tr}(x,z)$ = $U_0 e^{i[kz+\varphi(x)]}$ is tilted to approximately $\hat{\mathbf{x}}\frac{d\varphi(x)}{dx} + \hat{\mathbf{z}}k$, i.e., by a position-dependent angle $\theta(x) \approx \frac{1}{k}\frac{d\varphi(x)}{dx}$. If the transmitted light is detected with an array of angle-sensitive photodetectors, the photocurrent signals I_{ph} produced by different pixels at different x locations will therefore vary with the local phase gradient $\frac{d\varphi(x)}{dx}$ of the object. Although it has been proven in Chapter 2 that a symmetric response of $\mathcal{R}(\theta)$ at around normal incidence provides phase imaging capability, here we focus more on the ability to determine the direction of local phase gradient, by means of obtaining a monotonic and asymmetric dependence of the response versus the local incidence angle [see Figure 3.1(b)]. Especially when $\mathcal{R}$ is linearly proportional to θ within the range of the divergence angle, it becomes convenient to achieve phase reconstruction under a simple algorithm. It also follows from this discussion that the contrast of the resulting image is ultimately limited by the photodetector responsivity slope $d\mathcal{R}/d\theta$ in the limit of small θ. As a result, devices with an asymmetric angular response are preferable for this application compared to symmetric devices (where $d\mathcal{R}/d\theta$ vanishes for $\theta = 0$).

This type of directional image sensors is similar to the design introduced in the last chapter where the coupling peaks are contributed by the diffractive gratings, and its desired angular asymmetry is produced by an additional set of reflector metasurface. Devices based on a similar concept, with responsivity peaked at geometrically tunable angles over an ultrawide field of view of ~150°, have been

reported by a previous group member Leonard Kogos to enable flat lensless compound-eye vision. Alternative device configurations for angle-sensitive vision that have been demonstrated previously include the use of lenslet arrays (Duparré et al., 2004), stacked gratings based on the Talbot effect (Gill et al., 2011), and micro-apertures across adjacent pixels (Yi et al., 2021). For phase imaging applications, the key advantage of the configuration of Figure 3.2(a) is the ability to be designed with particularly sharp asymmetric responsivity peaks of large slope $d\mathcal{R}/d\theta$.

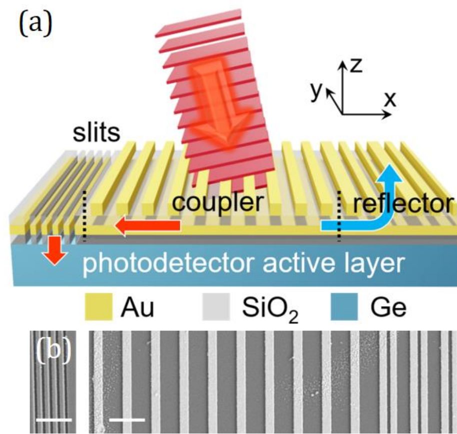

Figure 3.2: Asymmetric metasurface photodetectors. (a) Schematic device structure and principle of operation. (b) Top-view SEM images of an experimental sample, showing the slits (left image) and nanostripes (right). The scale bars are 2 μm. In this device, the Au stripes have a width of 440 nm. The grating coupler contains 10 lines with a period Λ = 1432nm.

To demonstrate the resulting wavefront sensing capabilities, here we have developed a tailor made device for this application, measured its angle-dependent responsivity, and then used the experimental data in conjunction with computational imaging techniques to evaluate the phase contrast images produced by full pixel arrays of these sensors. Results show that a minimum detectable phase contrast as

small as 8 mrad can be achieved, highlighting the promise of these angle-sensitive photodetectors to substantially miniaturize and simplify phase imaging systems while still providing state-of-the-art sensitivity.

The phase measurement carried out by these devices is conceptually similar to the DPC approach, in which a reciprocal-space asymmetry is introduced in the sample illumination, in the pupil plane (Kim et al., 2012), or by split detectors in a scanning microscope (Hamilton and Sheppard, 1984), to convert phase gradients into intensity variations. This approach has been employed for *quantitative* phase reconstruction by sequentially recording one or multiple pairs of DPC images with mirrored asymmetric illumination (Mehta and Sheppard, 2009), (Tian and Waller, 2015), (Bonati et al., 2020). The two images in each intensity pair are subtracted from each other to remove the unknown background, and the process of phase differentiation is then digitally inverted by a deconvolution algorithm. As shown in the following, the same idea can be implemented with an array of asymmetric angle-sensitive photodetectors where alternating pixels feature equal and opposite responsivity functions $\mathcal{R}_{+}(\theta) = \mathcal{R}_{-}(-\theta)$. With this configuration, the two mirrored DPC images required for background subtraction are acquired simultaneously (i.e., in a single shot) by the two types of pixels. Correspondingly, the overall measurement can be significantly simplified compared to previous quantitative DPC setups, because it does not require any specialized time-modulated directional sources (Mehta and Sheppard, 2009), (Tian and Waller, 2015), (Bonati et al., 2020) or beam scanning (Hamilton and Sheppard, 1984). As a result, this approach is particularly promising

for applications where space and time are highly constrained, such as point-of-care and in vivo microscopy, endoscopy, and imaging of freely moving objects.

3.2 Design and Simulations

In the device architecture of Figures 3.2(a) and 3.2(b), the illumination window of a photodetector is coated with a SiO_2/Au/SiO_2 stack, with the same design parameters as described in Chapter 2 Section 3. A periodic array of Au nanostripes (grating coupler) is then introduced over the top SiO_2 layer, surrounded on one side by a set of subwavelength slits (same design as before) perforated through the stack and on the other side by a short section of Au nanostripes of different widths (reflector). Backward traveling SPPs are guided down by the slits and become electronic signal. Forward traveling SPPs eventually arrive at the reflector, which is designed to scatter them back into radiation propagating away from the device into the air above. Briefly, the nanostripe widths in this reflector section are selected to produce a linear scattering phase profile for the incoming SPPs (and therefore suppress all diffraction channels except for the -1 order) based on the notion of gap-plasmon metasurfaces (Sun et al., 2012). Three nanostries with a center-to-center spacing of 508 nm combines into a section, and there are three identical sections of reflectors placed in a group with a separation of 1524 nm. The widths of the nanostripes are 0 (blank), 303 nm, and 340 nm, forming a reflected phase gradient pointing towards the grating. With this arrangement, all forward traveling SPPs can be scattered away from the device surface within the smallest possible area. Altogether, the composite metasurface comprising the metal film, grating, slits, and reflector therefore behaves like an angle-selective filter for the light transmitted into, and ultimately absorbed by, the photodetector. The required asymmetric angular response for quantitative phase

imaging is enabled by the aforementioned diverging action of the slits and reflector on oppositely traveling SPPs.

The specific device developed in this work features a narrow responsivity peak $\mathcal{R}(\theta)$ centered at $\theta \approx 2°$, i.e., only slightly offset from normal incidence to maximize the slope $d\mathcal{R}/d\theta$ at $\theta = 0$. The key geometrical parameters, listed in the caption of Figure 3.2, were optimized via FDTD simulations. Because of the diffractive nature of the device operating principle, the angular peak position is sensitive to the incident wavelength, and operation near $\lambda = 1550$ nm is considered throughout this work. The resulting phase imaging system is therefore primarily intended for monochromatic (i.e., laser light) illumination, although high spatial coherence is not needed [unlike typical interferometric setups, which correspondingly often suffer from speckle artifacts (Park et al., 2018), (Paganin and Nugent, 1998), (Mehta and Sheppard, 2009)]. The grating-coupler nanostripe width w and period Λ are 440 nm and 1432 nm, respectively, selected to produce efficient excitation of SPPs by light incident at the desired angle of peak detection (~2°) according to the Bragg condition. The number of nanostripes in the grating is 10, selected to minimize the angular width of the responsivity peak (based on the interplay between SPP propagation losses and diffraction effects), while at the same time maintaining a reasonably small pixel size (21.8 μm, including the slits and reflector section).

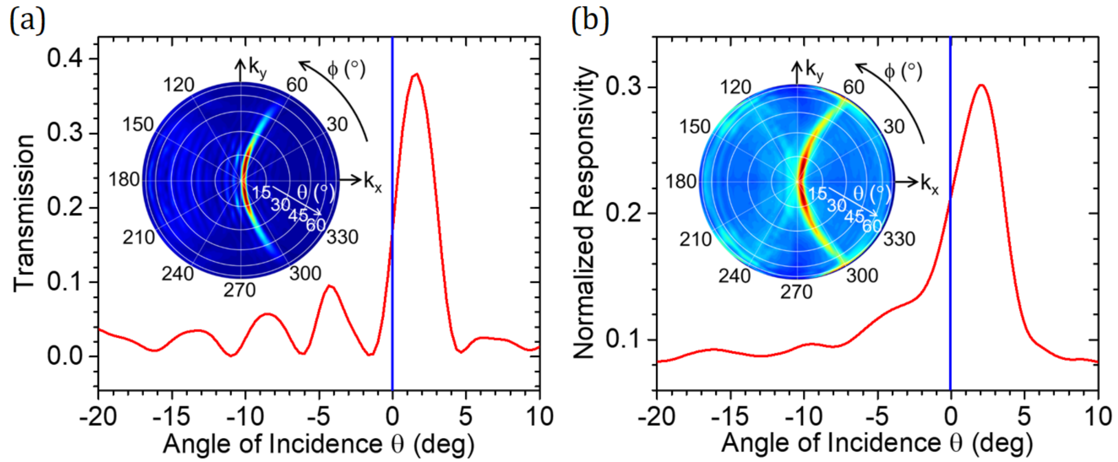

Figure 3.3: Performance of asymmetric metasurface photodetectors. (a) Inset: calculated transmission coefficient through the metasurface of this device for *p*-polarized incident light at λ = 1550 nm versus polar θ and azimuthal ϕ illumination angles. Main plot: horizontal line cut of the color map. (b) Inset: measured angular dependence of the responsivity of the same device, normalized to the normal-incidence responsivity of an identical photodetector without any metasurface. Main plot: horizontal line cut of the color map. The vertical blue lines in (a) and (b) indicate normal incidence.

Figure 3.3(a) presents simulation results for the p-polarized power transmission coefficient of the optimized metasurface as a function of polar θ and azimuthal ϕ angles of incidence. The figure inset shows the full angular response across the entire hemisphere, obtained from a three-dimensional FDTD simulation based on the principle of reciprocity. The result was generated by computing the far-field radiation pattern in the air above the device for an electric dipole source positioned in the device substrate below the slits. In this simulation, a three-dimensional computational domain is employed, with PMLs. All relevant materials are described by their complex permittivity from a built-in database in the FDTD software. By reciprocity (Potton, 2004), the calculated pattern is proportional to the local field

intensity at the dipole position produced by an incident plane wave as a function of illumination angles. This approach for computing the angular response of our devices is particularly convenient in terms of computational time, as all angles are covered in a single simulation. The main plot shows the horizontal line cut of the color map (i.e., transmission versus θ for $\phi = 0$). To calibrate the resulting color map, we have conducted additional two-dimensional simulations with Bloch boundary conditions on the lateral boundaries enclosing a full pixel. In these calculations, the metasurface is illuminated with a p-polarized plane wave and the transmitted light intensity into the device substrate is calculated for different values of the angle of incidence θ on the x-z plane. The results of these simulations are qualitatively in good agreement with the line cut of the main plot and allow calibrating its vertical axis to the metasurface transmission coefficient.

These simulation results reveal a narrow angular region of high transmission adjacent to normal incidence, with a characteristic C shape determined by the Bragg condition for the excitation of SPPs traveling along different directions. By design, the low-angle tail of the transmission peak is centered around $\theta = 0$ (vertical blue line in the main plot). The maximum transmission coefficient (at $\theta = 1.6°$) is over 38%, indicating that the transmission penalty introduced by the metasurface is reasonably small. Similar calculations for s-polarized incident light show negligible transmission at all angles, consistent with the longitudinal nature of SPP modes. As a result, these devices require polarized illumination for maximum detection efficiency.

3.3 Fabrication and Measurement results

If the metasurface just described is fabricated on the illumination window of an image sensor, the device responsivity can be expected to vary with angles of incidence exactly as in the color map of Figure 3.3(a), regardless of the photodetector operating principle. Fabrication process is the same as Section 2.4. Figure 3.3(b) shows SEM images of an experimental sample, highlighting the slits, grating, and reflector section. The completed device was characterized with angle-resolved photocurrent measurements under polarized laser light illumination, which can be found in Section 2.4. The incident wavelength λ was adjusted to optimize the position of the responsivity peak relative to normal incidence for maximum $d\mathcal{R}/d\theta$ at $\theta = 0$. All the experimental results presented below were measured with $\lambda = 1610$ nm, about 4% larger than the design value of 1550 nm. This rather small discrepancy is ascribed to similarly small deviations of the sample geometrical parameters from their target values (for example, the thickness of the SiO_2 spacer layer above the Au film, which affects the SPP effective index n_{SPP}).

With this adjustment, the measurement results are in good agreement with the design simulations. As shown in Figure 3.3(b), the measured responsivity peak is centered at 2.2° with a FWHM of 5.5°, reasonably close to the calculated values of 1.6° and 3.0°, respectively, from Figure 3.3(a). The vertical axis in Figure 3.3(b) is normalized to the responsivity of an otherwise identical reference sample without any metasurface. Correspondingly, a peak value of about 30% is obtained, again in reasonable agreement with the design simulations of the metasurface transmission.

The smaller peak height and larger FWHM observed in the experimental data likely originate from residual roughness in the Au film, which decreases the SPP propagation length and thus reduces the fraction of SPPs captured by the slits. The inset of Figure 3.3(b) also shows a weak signature of photocurrent measured through the excitation of forward traveling SPPs (faint C-shaped feature in the left half of the color map), which is attributed to a small misalignment of the slits relative to the grating section. However, this unintended photodetection channel does not significantly degrade the angular response asymmetry near normal incidence, as can be clearly seen in the line cut of the same figure.

3.4 DPC Imaging

3.4.1 Physical Model and DPC Imaging Simulation

Next, we consider an image sensor array based on the devices of Figure 3.3 and evaluate its phase contrast imaging capabilities. To that purpose, we employ the frequency-domain model developed in Chapter 2 to substantiate the use of similar plasmonic directional photodetectors for optical spatial filtering. The key conclusion of this model is that these devices sample the incident field distribution at their slit locations ($\mathbf{r} = \mathbf{r}_{\mathrm{sl}}^{\boldsymbol{n}}$ for the $\boldsymbol{n}^{\mathrm{th}}$ pixel in the sensor array), according to the coherent transfer function

$$\mathrm{t}(\mathbf{k}) \equiv \frac{\mathrm{E}_{\mathrm{SPP}}(\mathbf{k})}{\mathrm{E}_{\mathrm{in}}(\mathbf{k})} \propto \mathrm{e}^{-i\alpha \mathrm{k_x}}\sqrt{\mathcal{R}(\mathbf{k})}. \tag{3.1}$$

In this equation, $\mathbf{k} = (2\pi/\lambda)(\hat{\mathbf{x}}\cos\phi + \hat{\mathbf{y}}\sin\phi)\sin\theta$ is the in-plane wavevector (with $\hat{\mathbf{x}}$ perpendicular to the slits and nanostripes), $\mathrm{E}_{\mathrm{in}}(\mathbf{k})$ and $\mathrm{E}_{\mathrm{SPP}}(\mathbf{k})$ are the spatial Fourier transforms of the incident and SPP fields on the sensor array $\mathrm{E}_{\mathrm{in}}(\mathbf{r}_{\mathrm{sl}}^{\boldsymbol{n}})$ and $\mathrm{E}_{\mathrm{SPP}}(\mathbf{r}_{\mathrm{sl}}^{\boldsymbol{n}})$, respectively. Finally, the phase slope α is approximately equal to the distance between the slits and the pixel center, depending on the SPP propagation losses. The exact value of this parameter has actually no observable effect on the recorded images [by the shifting property of Fourier transforms, the phase factor $\mathrm{e}^{-i\alpha \mathrm{k_x}}$ in $\mathrm{t}(\mathbf{k})$ simply corresponds to a uniform displacement in real space by the amount α in the –x direction]. In the following, we use α = 8 μm, computed via FDTD simulations for the present device. For the responsivity function $\mathcal{R}(\mathbf{k})$ in Equation (3.1), we use the experimental data shown in the inset of Figure 3.3(b). Finally, $\mathrm{E}_{\mathrm{in}}(\mathbf{k})$

can be related to the Fourier transform of the object in the field of view $\mathrm{E}_{\mathrm{obj}}(\mathbf{k})$ according to $\mathrm{E}_{\mathrm{in}}(\mathbf{k}) = \mathrm{t}_{\mathrm{lens}}(\mathbf{k})\mathrm{E}_{\mathrm{obj}}(\mathbf{k})$, where $\mathrm{t}_{\mathrm{lens}}(\mathbf{k})$ is the pupil function of the optical imaging system. With these prescriptions, the photocurrent signal produced by each pixel, which is proportional to $|\mathrm{E}_{\mathrm{SPP}}(\mathbf{r}_{\mathrm{sl}}^{\boldsymbol{n}})|^2$, can be calculated for any given object as a function of pixel position $\mathbf{r}_{\mathrm{sl}}^{\boldsymbol{n}}$ across the array.

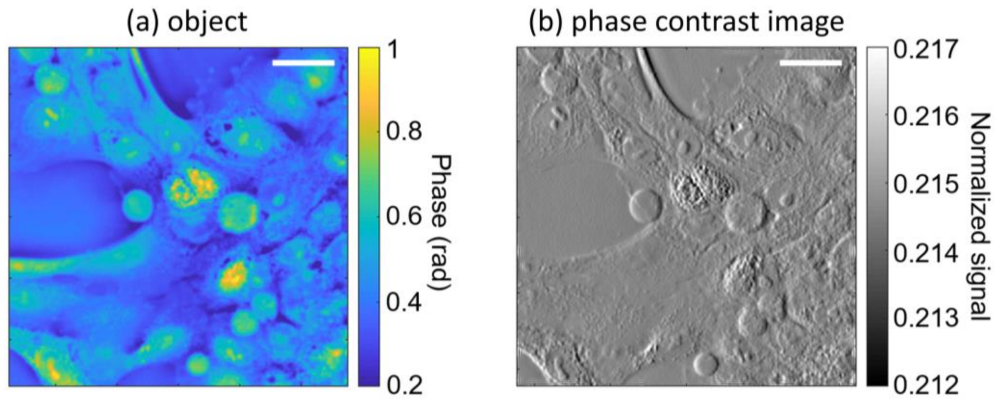

Figure 3.4: Computational phase contrast imaging results. (a) Representative phase object (MCF-10A cancer cells). (b) Corresponding image detected by an array of 512×512 angle-sensitive pixels modeled using the experimental data of Figure 3.3(c). The signal intensity in this plot is normalized to the photocurrent produced by an otherwise identical device without any metasurface under the same illumination conditions. The scale bars (referenced to the object space in both panels) are 50 μm.

As an illustration, we consider the phase object shown in Figure 3.4(a) [a sample of epithelial MCF-10A cancer cells, from (Tian et al., 2015)]. Using the method just described, we compute the corresponding image recorded by a sensor array consisting of 512×512 square pixels described by the responsivity data $\mathcal{R}(\mathbf{k})$ of Figure 3.3(b), combined with a telecentric 40× magnification system with NA = 0.8. Despite the transparent nature of the simulated object, a well resolved image is obtained, as shown in Figure 3.4(b). Specifically, the detected signals at the cell edges

are enhanced or decreased relative to the uniform background depending on the sign of the edge phase gradient along the horizontal (x) direction, in accordance with the asymmetric variation of $\mathcal{R}$ versus k_x around normal incidence. The resulting image contrast is therefore maximum for vertically oriented edges, and steadily decreases for edges oriented towards the horizontal direction. In the present approach, it could be eliminated by alternating pixels with orthogonally oriented nanostripes in a checkerboard pattern across the sensor array, as described in more details below. It should also be noted that the ~1550-nm operation wavelength of the present devices is not optimal for visualizing biological samples due to the background infrared water absorption. Nevertheless, the complex phase distribution of Figure 3.4(a) provides a particularly vivid illustration of the phase-imaging capabilities of these devices. The extension of the same device concept to visible wavelengths and broadband operation is addressed in the conclusion section.

3.4.2 Minimum Detectable Phase Contrast

In this section, I will also estimate the minimum detectable phase contrast with the metasurface of Figure 3.3. For that purpose, let's consider a simpler phase object consisting of y-oriented grating lines of variable contrast $\Delta\varphi$ [Figure 3.5(a)]. The phase slope at the line edges is taken to be as large as possible, but small enough to avoid any noticeable pixelation in the detected image. Figure 3.5(b) shows the resulting photocurrent signal I(x) as a function of pixel position, computed with the same procedure above and normalized to the photocurrent of identical uncoated photodetectors under the same illumination conditions. Following the work of (Bonati et al., 2020), the grating lines of Figure 3.5(a) can be regarded as detectable if the contrast-to-noise ratio of the image $\mathrm{CNR} = \frac{\Delta \mathrm{I}}{\mathrm{I_{bg}}}\mathrm{SNR}(\mathrm{I_{bg}})$ is larger than 1. Here, ΔI = $|I_{max(min)} - I_{bg}|$ is the image contrast, where $I_{max(min)}$ is the maximum (minimum) signal at the positive (negative) edges of the grating lines, and I_{bg} is the background signal away from the edges [see Figure 3.5(b)]. The parameter SNR(I_{bg}) is the signal-to-noise ratio at the background signal level, which depends on the photodetector characteristics. For this analysis, we consider high-performance image-sensor photodiodes, where the dominant noise mechanism is generally shot noise and therefore the SNR is proportional to the square root of the signal. For optimized near-infrared photodiodes of comparable dimensions as the present devices, a SNR at full well capacity SNR_{sat} = 71.3 dB (3,670×) can be achieved (Murata et al., 2020). Furthermore, in the envisioned imaging system the optical source can be selected so

that the pixels reach full capacity when illuminated at their angle of peak detection, where the photocurrent signal (again normalized to an identical uncoated device) is I_{peak} = 30% [from Figure 3.3(b)]. Based on all these considerations, SNR(I_{bg}) can be determined from the background signal I_{bg} according to $\mathrm{SNR}(I_{bg}) = \mathrm{SNR}_{sat}\sqrt{I_{bg}/I_{peak}}$.

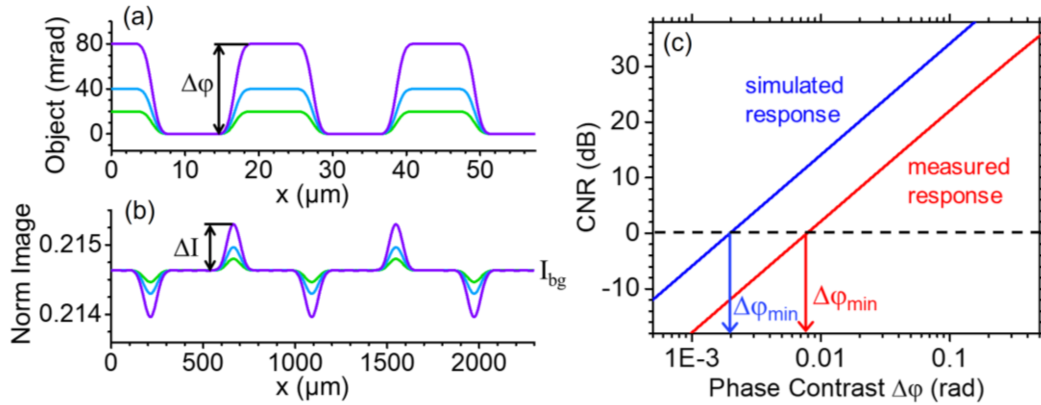

Figure 3.5. Minimum detectable phase contrast analysis. (a) Phase profiles of a one-dimensional grating for different values of the phase contrast Δφ. (b) Line cuts of the corresponding images detected by a 2D array of angle-sensitive pixels modeled using the experimental data of Figure 2(d), combined with a 40× magnification system with NA = 0.8. (c) Contrast-to-noise ratio (CNR) versus phase contrast Δφ for the object of (a), computed using the measured (red line) and calculated (blue line) angular response maps of the devices of Figure 2. The vertical arrows indicate the minimum detectable values of Δφ, below which CNR < 1.

Figure 3.5(c) shows the CNR computed with this model as a function of the object phase contrast Δφ, with the image (and therefore I_{max}, I_{min}, and I_{bg}) evaluated using the measured responsivity map $\mathcal{R}(\mathbf{k})$ of Figure 3.3(b) (red line) and the calculated map of Figure 3.3(a) (blue line). As indicated by the arrows in the same plot, the minimum detectable phase contrasts obtained from these traces are 8 mrad and 2

mrad, respectively. These values are on par with the sensitivity limits of standard DPC techniques (Bonati et al., 2020), which are based on more complex and bulkier setups as described above. Even smaller phase contrasts ($\lesssim$ 1 mrad) can be detected using interferometry (Ling et al., 2018) or a recently reported lock-in detection scheme (Bonati et al., 2022), at the expense however of a further increase in system and measurement complexity. The results plotted in Figure 3.5(c) therefore indicate that the present approach is fully suitable for high-sensitivity phase imaging applications, with the distinct advantage of enhanced miniaturization and portability. The comparison between the two traces in this figure also shows that, while the sensitivity is somewhat degraded by fabrication imperfections, state-of-the-art performance is still predicted for the experimental metasurfaces reported in this work, when combined with optimized image sensors.

3.4.3 Single-Shot Quantitative Phase Reconstruction

Our devices also naturally lend themselves to single-shot quantitative phase reconstruction, using the array configuration shown schematically in Figure 3.6(a). Here the array is partitioned into blocks of four adjacent pixels, each coated with the metasurface of Figure 3.3 oriented along one of four orthogonal directions. In the following discussion, each type of pixels will be labeled by the unit vector perpendicular to the metasurface nanostripes and pointing away from the slits ($\hat{\mathbf{u}} = \pm\hat{\mathbf{x}}$ or $\pm\,\hat{\mathbf{y}}$). The photocurrent signals $I_{\hat{\mathbf{u}}}(\mathbf{r})$ measured by all pixels of each type across the whole array as a function of pixel-block center position $\mathbf{r}$ provide an edge-enhanced image of the phase object [such as for example Figures 3.4(b) and 3.5(b) for $\hat{\mathbf{u}} = +\hat{\mathbf{x}}$]. In these images, each edge of the phase object transverse to the $\hat{\mathbf{u}}$ direction produces a peak or a dip (depending on the sign of the edge slope) over a constant background, which in turn is proportional to the incident optical power P and thus is generally unknown. Because of the asymmetric nature of these angle-sensitive devices, a peak over the background in $I_{+\hat{\mathbf{u}}}(\mathbf{r})$ corresponds to a dip in $I_{-\hat{\mathbf{u}}}(\mathbf{r})$ and vice versa. As a result, if the readout signals of the two pixels oriented along equal and opposite directions in each block are digitally normalized to their sum and subtracted from each other, the unknown background is subtracted out. The resulting signals

$$S_u(\mathbf{r}) \equiv \frac{I_{+\hat{\mathbf{u}}}(\mathbf{r}) - I_{-\hat{\mathbf{u}}}(\mathbf{r})}{I_{+\hat{\mathbf{u}}}(\mathbf{r}) + I_{-\hat{\mathbf{u}}}(\mathbf{r})} \tag{3.2}$$

(for u = x and y) can therefore be used for quantitative phase reconstruction.

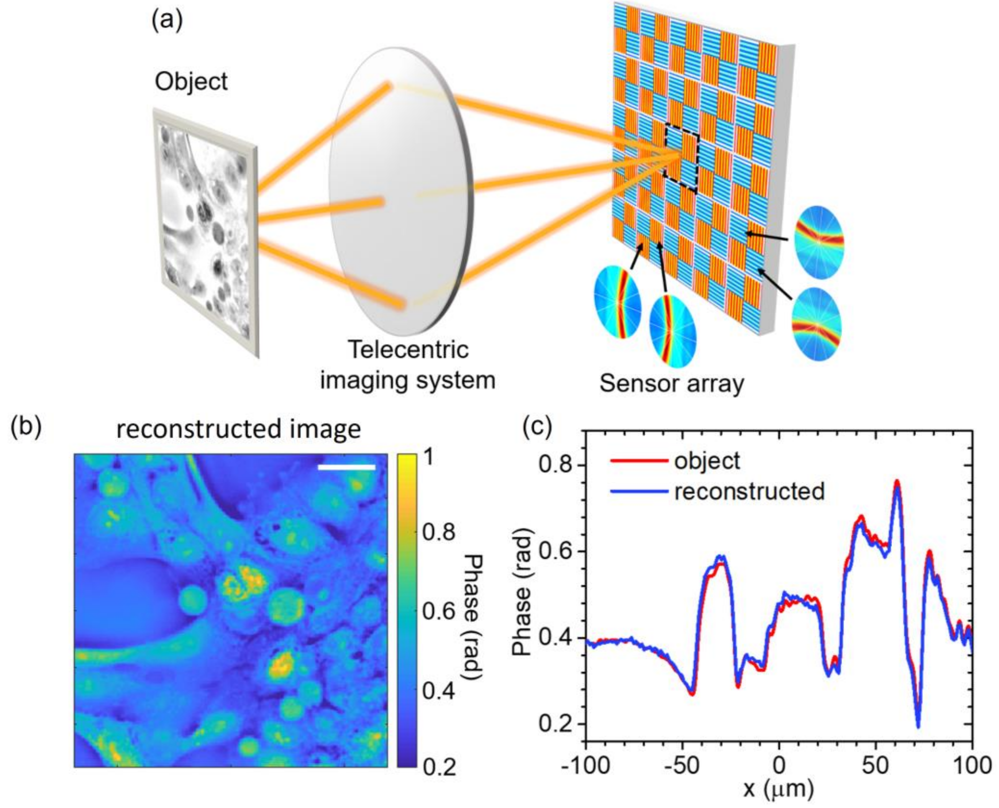

Figure 3.6: Computational quantitative phase imaging results. (a) Measurement protocol, where the sensor array is partitioned into blocks of four adjacent pixels coated with the metasurface of Figure 3.3 oriented along four orthogonal directions. One representative pixel block is indicated by the dashed lines. The experimental angular response maps of all four pixels in each block are also shown. (b) Reconstructed phase distribution of the MCF-10A cell sample of Figure 3.4(a). (c) Red trace: phase profile along the horizontal line at y = 0 of the same sample. Blue trace: reconstructed phase profile from (b).

In particular, for a pure phase object with sufficiently small phase $\varphi(\mathbf{r})$, the Fourier transforms of $S_u(\mathbf{r})$ and $\varphi(\mathbf{r})$ are linearly proportional to each other, i.e.,

$$S_u(\mathbf{k}) = H_u(\mathbf{k})\varphi(\mathbf{k}), \tag{3.3}$$

with transfer function (for u = x)

$$H_x(\mathbf{k}) = i t_{\text{lens}}(\mathbf{k})\left\{\sqrt{\frac{\mathcal{R}(\mathbf{k})}{\mathcal{R}(0)}} - \sqrt{\frac{\mathcal{R}(-\mathbf{k})}{\mathcal{R}(0)}}\right\}, \tag{3.4}$$

where i is the imaginary unit and $\mathcal{R}(\mathbf{k})$ is the responsivity map of Figure 3.3. For u = y, the same expression applies with $\mathcal{R}(\mathbf{k})$ rotated by 90°. The key role played by the asymmetric nature of our devices is clearly evidenced in Equation (3.4), where the transfer function $H_u(\mathbf{k})$ would be identically zero for a symmetric responsivity map subject to $\mathcal{R}(\mathbf{k}) = \mathcal{R}(-\mathbf{k})$. The derivation of Equations (3.3) and (3.4) is detailed in the Supplementary Material and builds on prior work on quantitative DPC imaging with time-modulated directional sources (Tian and Waller, 2015). A similar expression can also be derived for the more general case of an object that introduces both amplitude and phase modulation upon light transmission or reflection. Importantly, the transfer function of Equation (3.4) does not depend on the incident optical power P and is fully determined by intrinsic properties of the imaging optics [$t_{lens}(\mathbf{k})$] and of the image sensors [$\mathcal{R}_u(\mathbf{k})$, which can be measured in the initial device calibration as in Figure 3.3(b)]. Therefore, the phase profile $\varphi(\mathbf{r})$ can be retrieved quantitatively from the measured images $S_u(\mathbf{r})$ by inverting Equation (3.3). To avoid numerical artifacts associated with the zeros of the transfer function $H_u(\mathbf{k})$, we use the Tikhonov inversion method (Tian and Waller, 2015), whereby the reconstructed profile is

$$\varphi(\mathbf{r}) = \mathcal{F}^{-1}\left\{\frac{\sum_{u=x,y} H_u^*(\mathbf{k}) S_u(\mathbf{k})}{\sum_{u=x,y} |H_u(\mathbf{k})|^2 + \alpha_T}\right\}. \tag{3.5}$$

In this equation, $\mathcal{F}^{-1}\{\}$ indicates the inverse Fourier transform, α_T is a regularization parameter, and both $S_x(\mathbf{r})$ and $S_y(\mathbf{r})$ are used simultaneously to allow for isotropic phase reconstruction.

An illustration of this protocol is shown in Figures 3.6(b) and (c) for the phase object of Figure 3.4(a). Here the edge-enhanced images recorded by the four types of pixels in the sensor array, i.e., $I_{\pm\hat{\mathbf{x}}}(\mathbf{r})$ and $I_{\pm\hat{\mathbf{y}}}(\mathbf{r})$, were computed with the frequency-domain model of Equation (3.1), again using the experimental data of Figure 3.3(b) for the angle-dependent responsivity. Gaussian noise (with SNR estimated as described above) was then added to each image, and the results were used to evaluate the normalized signals $S_x(\mathbf{r})$ and $S_y(\mathbf{r})$ of Equation (3.2). In passing it should be noted that, in this normalization step, each peak and dip in the phase contrast image is automatically doubled in height, while the SNR is degraded by a factor of $\sqrt{2}$; as a result, the CNR is increased by $\sqrt{2}$, leading to a proportional decrease in the minimum detectable phase contrast. Given $S_x(\mathbf{r})$ and $S_y(\mathbf{r})$, Equation (3.5) was finally employed to reconstruct the phase profile $\varphi(\mathbf{x})$ of the MCF-10A-cell sample of Figure 3.4(a). The result, shown in Figure 3.6(b), is in excellent agreement with the original object. For a more direct quantitative comparison, the red and blue traces in Figure 3.6(c) show, respectively, the original and reconstructed phase profile along the horizontal line at $y = 0$ of the same sample. Only very small discrepancies are observed in this plot, which are attributed to the weak-phase-object approximation used in the derivation of Equations (3.3) and (3.4). A similar small-signal linear approximation is also used in standard DPC techniques for quantitative phase reconstruction, where multiple images of the object are recorded sequentially under different asymmetric illumination conditions (Mehta and Sheppard, 2009), (Tian and Waller, 2015), (Bonati et al., 2020).

Figure 3.7 shows additional phase contrast imaging results for the MCF-10A cell object of Figure 3.4(a), constructed with the procedure described above. The image of Figure 3.7(a) is computed for an array of uncoated pixels, whose responsivity is essentially constant with incident wavevector (other than the decrease in transmission through the illuminated surface at large angles described by the Fresnel formulas). As expected, the transparent cells could not be visualized with this arrangement, except for negligibly small diffraction fringes. Figures 3.7 (b) and (c) show the images computed with the experimental responsivity map of Figure 3.3(b), for an array of identical metasurface devices with the nanostripes oriented along the vertical and horizontal directions, respectively [Figure 3.7(b) is the same as Figure 3.4(b)]. Finally, Figure 3.7(d) shows the phase contrast image recorded with the array configuration of Figure 3.6(a), where pixels with metasurfaces oriented along orthogonal directions are combined in a checkerboard pattern.

The image sensors discussed in this chapter thus allow for similar results, but with a significantly smaller system footprint and simpler measurement protocol. Furthermore, in the present approach, all the required images are collected simultaneously by the different types of pixels, which is beneficial for the purpose of increasing the frame rate (at the expense, however, of a proportional decrease in spatial resolution).

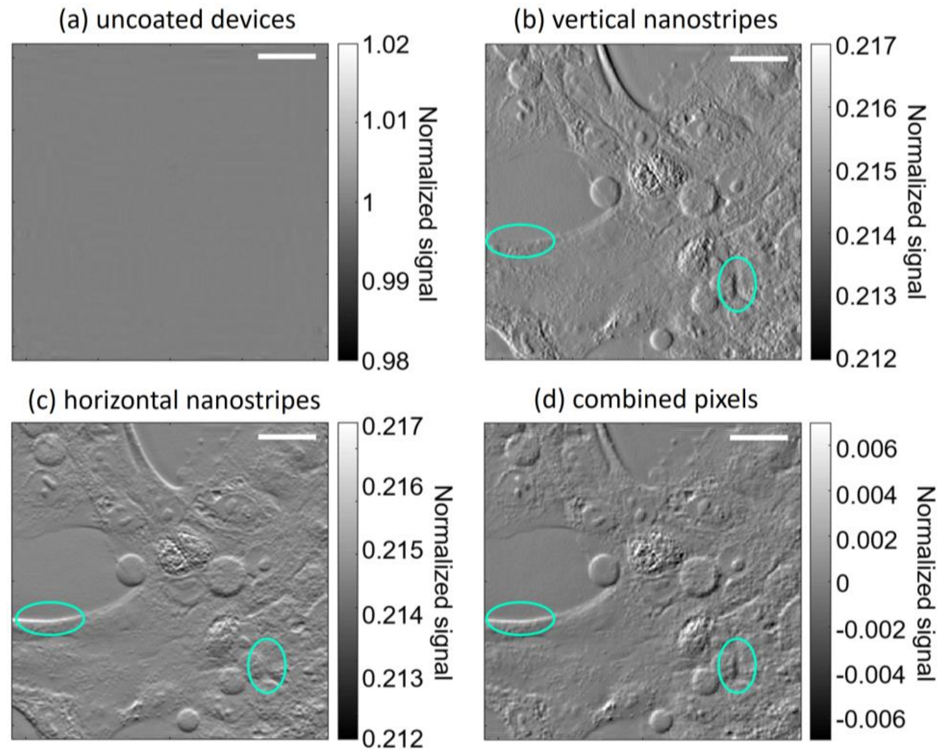

Figure 3.7: Additional phase contrast imaging results. (a) Image of the phase object MCF-10A cells of Figure 3.4(a) detected by an array of uncoated pixels. (b) Same as Figure 3.4(b). (c) Same as (b) but with the metasurface on each pixel rotated clockwise by 90°. (d) Image of the same object for an array partitioned into blocks of four adjacent pixels, each coated with the metasurface of Figure 3.3(b) oriented along one of four orthogonal directions. Specifically, this plot shows the signal $I_{+\hat{\mathbf{x}}}(\mathbf{r}) - I_{-\hat{\mathbf{x}}}(\mathbf{r}) + I_{-\hat{\mathbf{y}}}(\mathbf{r}) - I_{+\hat{\mathbf{y}}}(\mathbf{r})$. In all four panels, the simulated array consists of 512×512 pixels and is combined with a 40× magnification system with NA = 0.8. The green ovals in (b), (c), and (d) highlight two features of the phase objects that are not clearly resolved depending on the metasurface orientation. The scale bars are 50 mm in all panels.

3.5 Conclusion

In this chapter, I have introduced my study of a new type of image sensors that allow for the direct visualization of transparent phase objects with a standard camera or microscope configuration. Different from Chapter 2, the key innovation of these devices is a metasurface coating that creates an *asymmetric* dependence of responsivity on illumination angle around normal incidence. This arrangement produces a high sensitivity to wavefront distortions caused by light propagation through a phase object, with state-of-the-art minimum detectable phase contrasts below 10 mrad. At the same time, the combination of pixels with equal and opposite angular response can be employed to normalize out the unknown incident power, and thus perform quantitative phase reconstruction in a single shot. The specific devices developed in this work rely on a metallic metasurface design suitable for operation at near-infrared wavelengths, where plasmonic absorption losses are quite small. The same idea could also be extended to visible-range operation, e.g., by replacing SPPs with dielectric waveguide modes and the Au nanostripes with dielectric nanoparticles arranged in a gradient-metasurface architecture to introduce the required asymmetry. Similar configurations could also be designed to further tailor the angular response, including for example isotropic or vortex-like shapes, and to produce broadband achromatic operation by metasurface dispersion engineering.

CHAPTER FOUR

Gradient-Metasurface Directional Photodetectors

4.1 Introduction

Directional image sensors are focused on throughout this Dissertation. The simplest implementation of such sensors involves paired combinations of suitably aligned microlenses and photodetectors (Lam, 2015), (Duparré et al., 2005). However, the existing systems suffer from limited spatial resolution and field of view, due to a fundamental tradeoff between size and numerical aperture in microlenses. An alternative approach involves the integration of diffractive elements on the photodetector illumination window to selectively transmit or block incident light depending on its direction of propagation. In one implementation, stacked gratings based on the Talbot effect have been used to produce a sinusoidal dependence of responsivity on angle of incidence (Jayasuriya et al., 2015), (Chen et al.,, 2016). In my previous works described in Chapter 2 and 3, we have developed unique photodetectors coated with plasmonic grating couplers featuring geometrically tunable peaks in their angular response. Several advanced imaging functionalities have also been demonstrated in the literature, including microscale cameras and lightfield imagers (Jayasuriya et al., 2015), lensless compound-eye vision (Kogos et al., 2020), optical edge detection, and quantitative phase imaging. However, the diffractive operating principle of these devices severely restricts their design flexibility, for example to decouple the angular and spectral responses for broadband achromatic operation. Wavefront sensing has also been reported based on near-field

scattering by micro-apertures across adjacent pixels (Yi et al., 2021), which provides a rather weak wavelength dependence at the expense of limited ability to control the shape of the angular response.

In this chapter, I will describe the development of angle-sensitive photodetectors based on the GMS design platform, motivated by its distinctive ability to provide enhanced miniaturization and multifunctional operation. To date, plasmonic and dielectric GMSs have mostly been used as free-space passive components, designed to shape the incident optical wavefronts according to a desired device operation (e.g., light focusing, polarization control, holographic projection) (Genevet et al., 2017). In the context of integrated devices, they have been employed for the detection of orbital angular momentum (Genevet et al., 2012), directional light emission (Wang et al., 2020), (Iyer et al., 2020), (Xie et al., 2020), and to interface waveguided modes to free space radiation (Lingling et al., 2013), (Pors et al., 2014), (Meng et al., 2021), (Huang et al., 2023), (Chen et al., 2023). GMS is a type of metasurfaces whose local response to the incident light forms a gradient profile along certain direction of the surface plane. For the specific example here, our GMS generates an in-plane phase gradient of the near-field reflected light close to the interface, which results in a controlled addition to the in-plane momentum of the photons.

4.3 Design and Simulations

My designed device structure is shown schematically in Figure 4.1(a). The photodetector illumination window is coated with a SiO_2/Au/SiO_2 stack supporting an aperiodic array of 50-nm-thick Au nanostripes. Apart from the nanostripes' dimensions in the x-y plane, all other design parameters are inherited from the last two chapters. In this geometry, the p-polarized reflection phase of each unit cell can be tuned across a large fraction of the entire 2π phase space by varying the nanostripe width L_x, as shown by the FDTD simulation results of Figure 4.1(b) for an incident wavelength λ_0 = 1550 nm. This large phase tunability is enabled by the coupling between the nanostripe plasmonic resonances and their dipolar images in the metal film (Pors et al., 2014), (Chen et al., 2023), the same as how the reflectors work in Chapter 3. At the same time, the unit-cell reflection coefficient remains quite large (> 90%) for all values of L_x [blue trace of Figure 4.1(b)]. In the present devices, the nanostripe widths are selected so that their reflection phase varies linearly with their center position along the x direction, leading to a discretized version of the linear phase profile $\Phi_x = \xi x$ (mod 2π) [Figure 4.1(c)]. As a result, SPPs on the underlying metal film can be excited by p-polarized light with in-plane wavevector $\mathbf{k}_{||}$ satisfying the phase-matching condition $\mathbf{k}_{||} + \xi\hat{\mathbf{x}} = \mathbf{k}_{SPP}$. These SPPs are then intercepted by the adjacent slits where they are preferentially scattered into the Ge substrate, due to its higher refractive index (and therefore higher density of radiation modes) compared to the air above. Light incident along any other direction is instead reflected back into free space, leading to a highly directional asymmetric angular response.

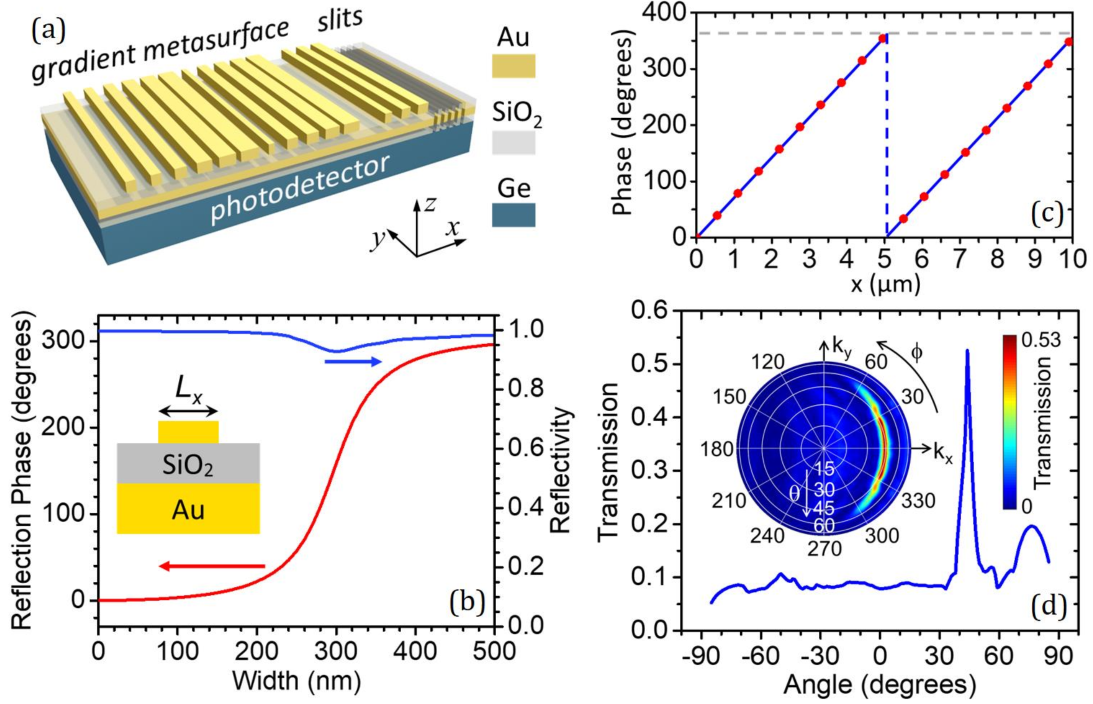

Figure 4.1: GMS design. (a) Schematic device geometry. (b) Reflection phase (red trace) and amplitude squared (blue trace) of the GMS unit cell shown in the inset, computed as a function of nanostripe width L_x for normally-incident x-polarized light at λ_0 = 1550 nm. (c) Reflection phase vs nanostripe center position for a GMS designed to promote directional photodetection peaked at $\theta_p = 45°$. (d) Transmission coefficient of the GMS of (c) computed as a function of polar illumination angle θ on the x-z plane. Inset: transmission coefficient of the same device vs in-plane wavevector of the incident light, computed by the principle of reciprocity (see Section 3.2).

Figure 4.1(d) shows FDTD simulation results for the transmission coefficient of a GMS designed for peak detection at a polar angle $\theta_p = 45°$, computed as a function of polar θ and azimuthal ϕ illumination angles, or equivalently as a function of $\mathbf{k}_{||}$. The array contains 29 unit cells, where neighboring nanostripes have center-to-center distance δx = 550 nm and reflection phase differing by $\delta\Phi_x = 39°$. The nanostripe widths range from 0 (corresponding to a missing nanostripe in the respective unit cell) to 500 nm. The resulting phase gradient

$\xi = \delta\Phi_x/\delta x$ produces the desired value of θ_p according to the phase matching condition for light propagating on the x-z plane, i.e., $2\pi\sin\theta_p/\lambda_0 + \xi = 2\pi n_{SPP}/\lambda_0$, where n_{SPP} is the SPP effective index. The design simulation results of Figure 4.1(d) are fully consistent with expectations. In the angular response map shown in the inset, high transmission through the GMS is obtained only for a narrow set of values of $\mathbf{k}_{||}$ determined by the SPP phase matching condition. For incident light propagating on the x-z plane [main plot of Figure 4.1(d)], the transmission peak is centered near the target detection angle $\theta_p = 45°$ with a small linewidth $\delta\theta = 6.0°$ FWHM and maximum value $T_{max} = 53\%$. For comparison, the Fresnel transmission coefficient of the Ge/air interface is about $T_{ref} = 62\%$, so that the transmission penalty introduced by the GMS is quite small ($T_{max}/T_{ref} = 85\%$). Additional simulation results (not included) show that for s-polarized light the GMS transmission is negligibly small (< 0.3%) at all angles, which is a consequence of the longitudinal nature of SPPs. This polarization dependence limits the device sensitivity for applications involving unpolarized light, although it could also be exploited to enable polarization vision for improved imaging contrast.

The nanostripe array just described essentially behaves like a plasmonic-waveguide coupler. Unlike typical applications of such devices, however, here the angular response must be controlled for all possible angles of incidence across the full hemisphere, which introduces two additional design considerations. First, the nanostripe dimensions must be selected to avoid any periodic repetition across the array. Otherwise, the GMS would simultaneously approximate discretized versions of additional linear phase profiles (modulo

2π) with slopes equal to integral multiples of ξ (Wang et al., 2020), leading to the appearance of additional peaks in the angular response. Second, it has been shown that a GMS with linear phase profile supports driven surface waves whenever no radiative channels exist for the reflected light (Sun et al., 2012). Such surface waves are also excited in the device of Figure 4.1 and can potentially produce a strong increase in the transmission background at large angles of incidence. In the present work, however, we found that, unlike SPPs, these surface waves are predominantly scattered into free space before reaching the slits, even for moderate separations (960 nm in Figure 4.1) between the slit and nanostripe sections. As a result, their contribution to the GMS transmission is quite small [e.g., limited to the small bump observed near 75° in Figure 4.1(d)].

The overall device architecture of Figure 4.1 was designed building on our prior work of Chapter 2 and 3, where directional photodetection was achieved through the selective excitation of SPPs by a periodic diffraction grating. Asymmetry in the angular response was then introduced by surrounding the grating coupler with slits on one side and a nanostripe-array "reflector" on the other, designed to scatter all incoming SPPs back into free space. Compared to these diffractive devices, the present approach has several important advantages. First, by removing the reflector section, the pixel size is reduced (without sacrificing angular sensitivity), leading to a proportional increase in spatial resolution for imaging applications. In fact, for the device of Figure 4.1(d), the full lateral dimension including GMS and slits section is 20.1 μm, versus 27.2 μm for the design reported in (Kogos et al., 2020) featuring the same angle of peak detection $\theta_p = 45°$ and similar (actually larger) linewidth $\delta\theta = 7.6°$.

Second, the present device also provides larger peak transmission [T_{max} = 53% in Figure 4.1(d) versus 38% for the same baseline structure of Chapter 3], and thus higher sensitivity. This improvement is mostly ascribed to the suppressed radiative scattering of SPPs in the presence of the GMS linear phase gradient compared to periodic gratings, leading to higher SPP coupling efficiency (Sun et al., 2012). Furthermore, with the use of more complex meta-unit geometries, the GMS platform of Figure 4.1 can be extended to enable capabilities that are not accessible with diffractive devices. Examples of practical interest include broadband achromatic operation, polarization independent response, and multifunctional operation such as the simultaneous measurement of angle of incidence and polarization.

4.3 Experimental Results

The GMS of Figure 4.1 was fabricated on the illumination window of a Ge metal-semiconductor-metal photoconductor. SEM images of a resulting sample are shown in Figures 2.

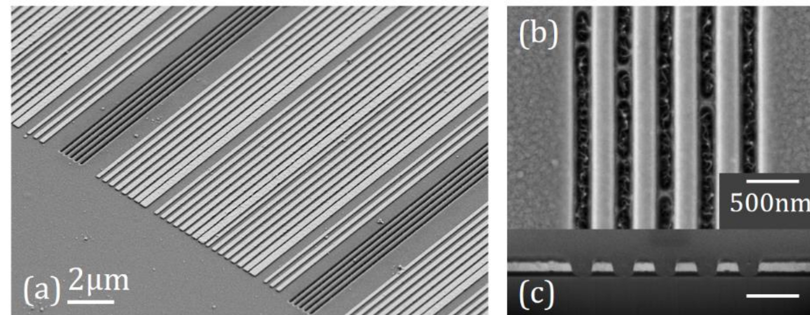

Figure 4.2: SEM images. (a) SEM image of the GMS. (b), (c) Top-view (b) and cross-sectional (c) SEM images of a slit section.

The fabrication process is identical to that introduced in Section 2.4 except the slits. In here, focused ion beam milling is used to pattern the slits after the completion of the GMS, following the recipe to be included in Section 5.2. It should also be noted that our experimental samples consist of a few (7-9) identical repetitions of the structure of Figure 4.1(a), with each nanostripe-array section surrounded symmetrically by two sets of slits [see Figure 4.2(a)]. This allows for a more conclusive demonstration of the asymmetric angular response of the GMS, where no SPPs are excited in the "backward" direction [the negative x direction of Figure 4.1(a)] which would otherwise cause crosstalk and spurious signals in a pixel array.

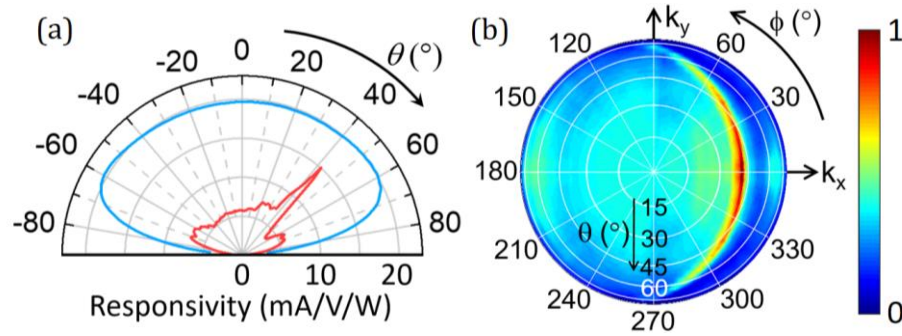

Figure 4.3: Measurement results for a GMS device based on the design of Figure 4.1 (a) Photocurrent measured with the same device (red trace) and with a nominally identical photodetector without any metasurface (blue trace) versus polar angle of incidence θ on the x-z plane of Figure 4.1(a). (b) Measured photocurrent of the same GMS device as a function of both illumination angles θ and ϕ. The device bias in these measurements is 1 V. The incident light is linearly polarized on the x-z plane with 1-mW power at the laser output and approximately 14 μW on the device surface.

The experimental samples were characterized by angle-resolved photocurrent measurements with polarized laser light, using the procedure described in Section 2.4. The incident wavelength was set to 1610 nm, selected to maximize the angular detection peak (similar results were measured over a small wavelength range of a few 10 nm, beyond which larger variations were observed). This optimal value is somewhat larger (~ 4%) than the design wavelength λ_0 = 1550 nm, which is attributed to small deviations of the device geometrical parameters from their target values. With this adjustment, the measurement results are in good agreement with the design simulations. A sharp peak centered at θ_p = 42° with FWHM of about 9° is observed in the plot of photocurrent versus polar angle of incidence on the plane perpendicular to the nanostripes [red trace in Figure 4.3(a)]. The angular response map across the full hemisphere [Figure 4.3(b)] contains a similarly narrow region of high responsivity with a characteristic curved shape that originates from the phase-

matched excitation of SPPs with different wavevectors. The peak-to-average-background ratio in this plot is ~3 versus ~5 in the GMS transmission map of Figure 4.1(d). This difference is mostly ascribed to the finite Au surface roughness in the experimental sample, which can scatter some of the incident light into SPPs regardless of its direction of propagation. Finally, the blue trace in Figure 4.3(a) was measured from a reference sample without any metasurface, but otherwise based on the same geometry and featuring similar dark resistance of ~800 Ω (in these Ge MSM devices the resistance has been observed to scale with quantum efficiency, likely due to fabrication-induced surface defects affecting the carrier density). At the angle of peak detection, the responsivity of the metasurface device (red trace) is as large as 75% of that of the reference sample, reasonably close to the predicted transmission ratio T_{max}/T_{ref} = 85% discussed above.

Finally, Figure 4.4 shows angle-resolved photocurrent data measured with two other devices based on the architecture of Figure 4.1(a), illustrating the geometrical tunability of the angle of peak detection. Both samples contain 31 nanostripes with nearest-neighbor separation and reflection-phase difference δx = 610 nm and $\delta\Phi_x$ = 74° in Figure 4.4(a), δx = 590 nm and $\delta\Phi_x$ = 20° in Figure 4.4(b). A pronounced peak in the angular response is again obtained in each plot, with maximum response at θ_p = 30° in Figure 4.4(a) and 60° in Figure 4.4(b), in good agreement with the SPP phase matching condition discussed above. The device of Figure 4.4(b) also features particularly large contrast and flat background, which may be related to its smaller

phase difference $\delta\Phi_x$ between neighboring nanostripes. In this case, the desired linear phase profile is more accurately approximated by the discretized version implemented in the GMS, leading to increased robustness against fabrication imperfections. Even sharper peaks can be expected by further optimizing the fabrication process to minimize surface roughness and improve control of the nanostripe dimensions.

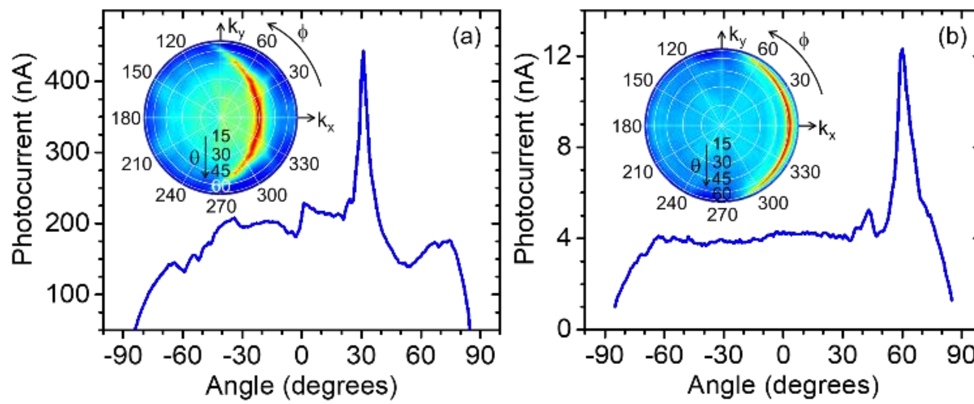

Figure 4.4: Geometrical tunability of the angle of peak detection θ_p. (a) Inset: photocurrent versus polar and azimuthal illumination angles measured with a GMS device designed for peak detection at θ_p = 30°. Main plot: horizontal line cut of the color map. (b) Same as (a) for a different device with θ_p = 60°. The incident wavelength is 1560 nm in (a) and 1630 nm in (b). The device of panel (a) [panel (b)] produces significantly larger [smaller] photocurrent compared to the sample of Figure 4.2, which correlates with its larger (smaller) dark resistance of over 1000 Ω (less than 20 Ω).

4.4 Conclusion

In summary, the research in this chapter demonstrate the ability to tailor the angular response of generic planar photodetectors through the integration of plasmonic GMSs on their illumination window and offer general guidelines for the design of suitable metasurfaces. The resulting directional image sensors are promising for multiple applications within the growing field of computational imaging. In analogy with extensive prior work in free-space flat optics, we also expect that more complex GMSs within the same architecture can be developed to enable additional advanced capabilities such as achromatic response, polarization independent operation, and multifunctional behavior.

CHAPTER FIVE

Wafer-Scale Upgrade of Sample Fabrication Process

5.1 Purpose of the Upgrade

In the work described so far in this Dissertation, our fabrication process for the MDM plasmonic photodetectors was based on a chip-by-chip style, which means that a clean germanium wafer is first diced into square pieces that are easy to handle, and each piece is then patterned with a single device. This approach is straightforward and very flexible for the initial device development.

At the same time, however, this method has several drawbacks. First, the fabrication efficiency and throughput are extremely limited, because each chip must be processed with many different tools, some of which (including spin coater, mask aligner, and EBL) are capable of wafer-scale operations. Second, some manual techniques such as photoresist development and BOE etching require precise control of time, so that the individual samples must be carefully handled one by one. Third, the most widely used photo-exposure tool MA-6 mask aligner is designed for operation with a whole wafer or with a small chip placed at the center of the sample loader. Performing photolithography chip-by-chip therefore requires either multiple photomasks with each pattern at the center, or a single mask with the chip placed at different locations in different exposures (Figure 5.1). These solutions are either cost ineffective or complicated, so an upgrade to whole-wafer processing is desired and it is expected to dramatically decrease the labor in batch production of the chips.

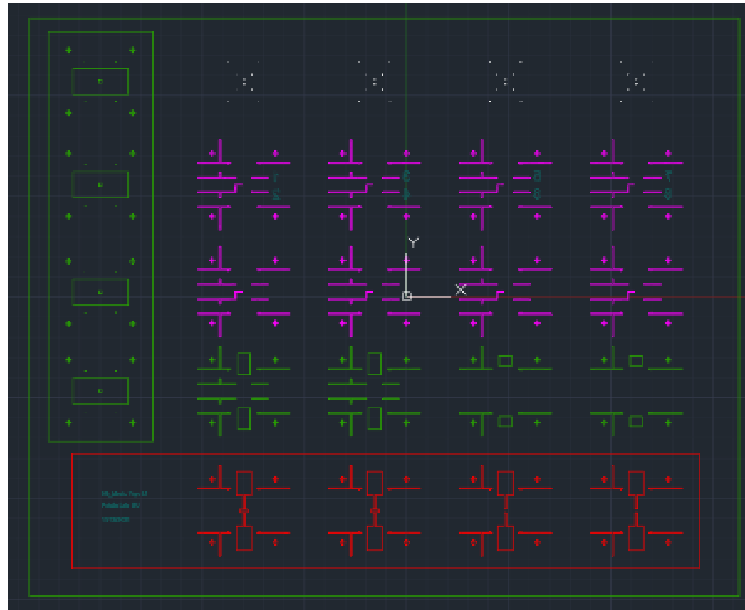

Figure 5.1: A photomask design with photolithography patterns at different positions.

Our previously completed samples also have large variations in performance and appearance, which can be attributed to the inconsistency of the processing conditions. By performing all the major fabrication steps in a wafer scale (including photolithography and EBL, both of which involve development, spin coating, baking, and exposure), significant improvement in uniformity can be achieved.

Furthermore, the relatively long chip-by-chip fabrication process often results in damaged substrates because germanium is fragile. The resulting germanium fragments become one of the main sources of contamination and damage on the surface of the samples. Therefore, it is highly desirable to minimize handling of the samples in order to increase yield. By processing at the wafer scale, the sample handling can also reduce the chances of creating fragments.

Another motivation for this upgrade is the use of more advanced EBL systems for

the fabrication of metasurfaces consisting of 2-dimensional nanoparticle arrays, which are promising to further extend the capabilities of the devices developed in this work. Although the EBL tool used so far (the Zeiss Supra40 VP equipped with NGPS) is nominally capable of patterning at sub-20 nm resolution, it is not paired with any dose and working-distance correction tools, and it is not able to perform actual aligned patterning. Therefore, beam calibration and exposure need to be done chip-by-chip, limiting the patterning speed and consistency among devices. The Elionix HS-50 at the Harvard Center for Nanoscale systems (CNS) provides not only higher resolution due to its 50-keV electron beams, but also wafer-scale aligned patterning with real-time working distance stabilization. In addition, the software is compatible with Beamer (see Figure 5.2), a very useful tool to simulate and compensate proximity effect caused by electron scattering, so that complex structures are easier to make. With the help of such modern EBL system, more than 20 metasurfaces for our devices can be exposed within 10 minutes.

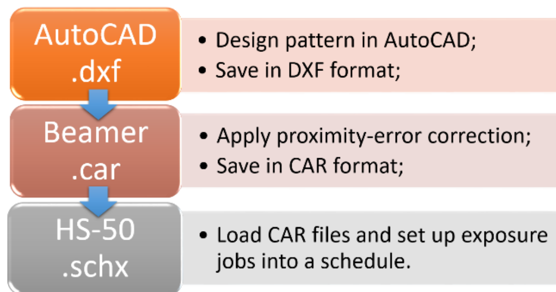

Figure 5.2: EBL patterning workflow. All three steps are capable of real-time visualization, including colored dose maps after applying proximity-error correction in Beamer. A schedule file contains all the information for the EBL machine to perform an exposure job.

However, it is worth noting that whole-wafer processing bears an intrinsic shortcoming: higher risk. If one step goes wrong, all the chips on the wafer are potentially affected. Therefore, this upgrade is supported with sufficient confidence in the reproducibility of each fabrication step.

5.2 Sample Fabrication Process

5.2.1 Wafer Preparation

The substrate (active material) of the studied devices is germanium, and thus we start the device fabrication process from a clean Ge wafer. The specifications are as follows:

a. Size: two-inch diameter and 0.5-mm thickness;

b. Supplier: University Wafers;

c. Doping: undoped;

d. Lattice orientation: (100);

e. Polish: double-sided polished.

The (100) lattice orientation provides the best adhesion for the Ti and SiO_2 layers to be deposited later and also ensures good quality of the subsequent dicing of the wafer. Before the first use of a wafer, it should go through the following cleaning:

a. At room temperature, put the wafer in a container filled with acetone and sonicate in an ultra-sonic bath for 1 minute;

b. Quickly transfer the wafer to another container filled with IPA before acetone on the wafer surface dries out, and sonicate for 1 minute;

c. Quickly transfer it to a third container filled with DI water before IPA dries out, and sonicate for 1 minute;

d. Remove the wafer from the DI water and quickly dry it with a nitrogen gun;

e. Bake it on a hot plate at 110 degrees for 2 minutes, and let it cool down.

These operations should remove all dust particles on the wafer surface, clean out the oxidization layer on the germanium surface (which is important for the adhesion and

conductivity between the Au contacts and the substrate) (Onsia et al., 2005), and evacuate all water molecules inside the wafer.

During the fabrication process, whenever the wafer is taken out of the cleanroom for longer than one day when contamination may happen, it is recommended to perform a gentle cleaning before the next step. If there are fragile thin films already deposited on the substrate, we should clean it without sonication:

a. At room temperature, put the wafer in a container filled with acetone and shake the container for 1 minute;
b. Quickly transfer it to another container filled with IPA before acetone on the wafer surface dries out, and shake it for 1 minute;
c. Remove the wafer from the IPA and quickly dry it using a nitrogen gun.

In this scenario, water treatment is undesired based on two considerations. First, water can dissolve the oxides on the germanium substrate, reducing the adhesion of the deposited thin films. Second, when there are SiO_2 layers, water can be absorbed inside and cause problems in the following steps. For example, if metal films are deposited on top of a SiO_2 layer with water, heating at higher than 100 degrees will release the water vapor and inflate bubbles between the layers, eventually breaking the metal film.

5.2.2 Process Flow

The device fabrication process is listed below, and Figure 5.3 shows a cross section of the complete layered structure (excluding alignment markers and other supporting features):

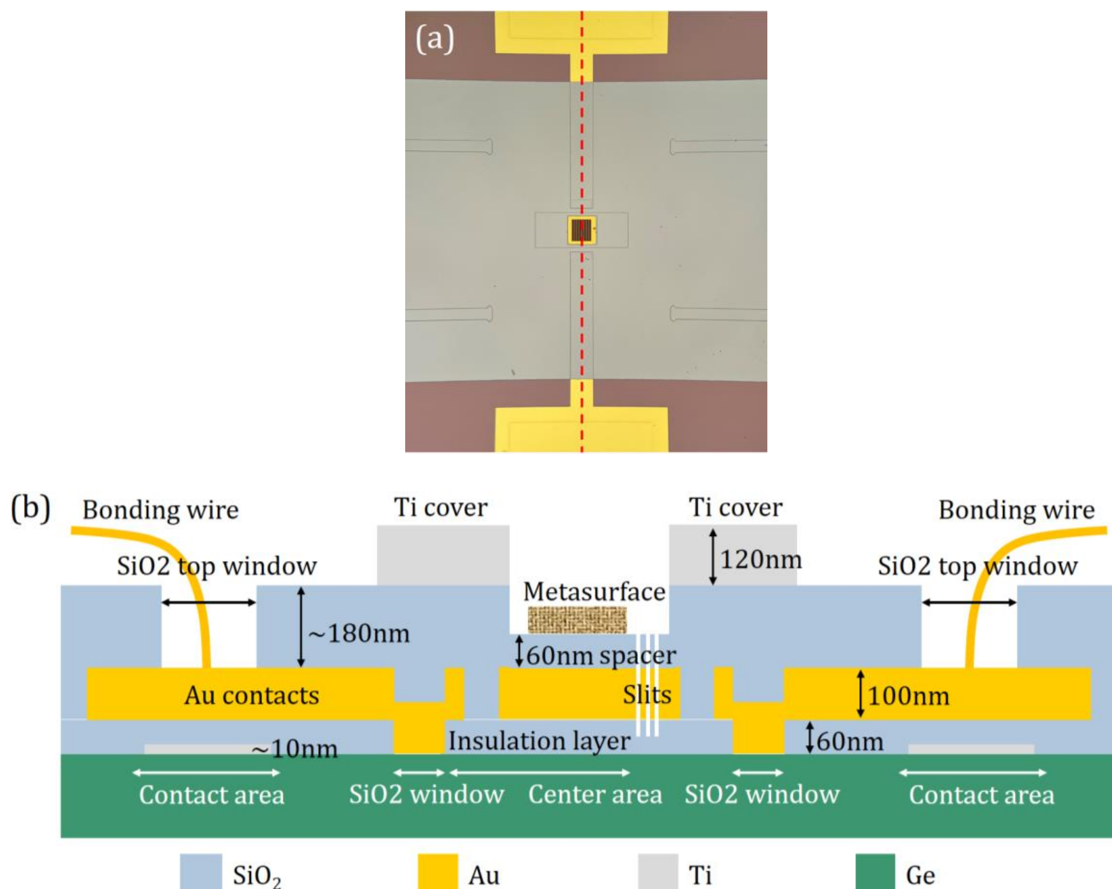

Figure 5.3: Structure of a completed device. (a) Top-view photograph of a completed device with alignment markers located at the four corners. (b) Schematic cross section along the red dashed line in (a). Dimensions are not to scale in order to show the relative positions of the different features.

1) Titanium pads.

After the wafer is cleaned as instructed in section 5.2.1, the first layer to be deposited and patterned are two titanium pads for each device. The purpose of

these Ti pads is to ensure strong SiO_2 adhesion to the Ge substrate during ball bonding in the final step. Photolithography with positive resist is used in this step, followed by e-beam evaporation of titanium and lift-off.

2) Bottom insulation layer.

An insulation layer of 60-nm-thick amorphous silica (SiO_2) is deposited by PECVD. The function of this dielectric layer is to isolate the metal contacts (to be made in the next step) from the substrate to reduce the background DC signal, and to prevent substrate doping by the metals. Next, photolithography and BOE etch are performed to open two small windows in the isolation layer in order to allow for current flow in and out of the substrate through the metal contacts.

3) Au layer and contacts.

The reflective gold layer of our metasurfaces needs to be sufficiently thick and to cover a large enough area to block any incident light from directly reaching the substrate (in addition to supporting the SPP modes in the directional photodetection process). Gold is also an ideal material for the two metal contacts connected to the substrate on both sides of the device area (and thus this layer should be thick enough to reduce the contact resistance). Therefore, a thickness of 100 nm is decided, and the Au film is deposited and patterned by positive-tone photolithography, e-beam evaporation, and lift-off. An adhesion layer of 5-nm titanium is also deposited under the gold. Additionally, plasma ashing is applied after the photolithography development step, in order to completely remove any photoresist residue for better film adhesion and flatness.

4) Top insulation and spacer layers

Due to the metallic sidewalls created in step 3, the titanium cover to be patterned in step 6 may touch the two arms of the Au contacts, resulting in a short circuit that would prevent proper device function. Thus, it is important to coat the Au reflective layer and contacts with a conformal insulator layer. First, a 120-nm-thick silica film covering the entire wafer is deposited by PECVD, followed by photolithography and BOE etch to uncover a square area at the center of each chip, as the designated area for the metasurface ("device area"). After the photoresist is removed, an additional 60-nm-thick SiO_2 spacer layer is deposited by PECVD with the same recipe as in step 2.

5) Metasurfaces

To pattern the metasurfaces, positive-tone aligned EBL is performed with detailed recipe in Section 5.2.3. Au nanoparticles of 50-nm height with an additional 5-nm Ti adhesion layer are then created by e-beam evaporation and lift-off.

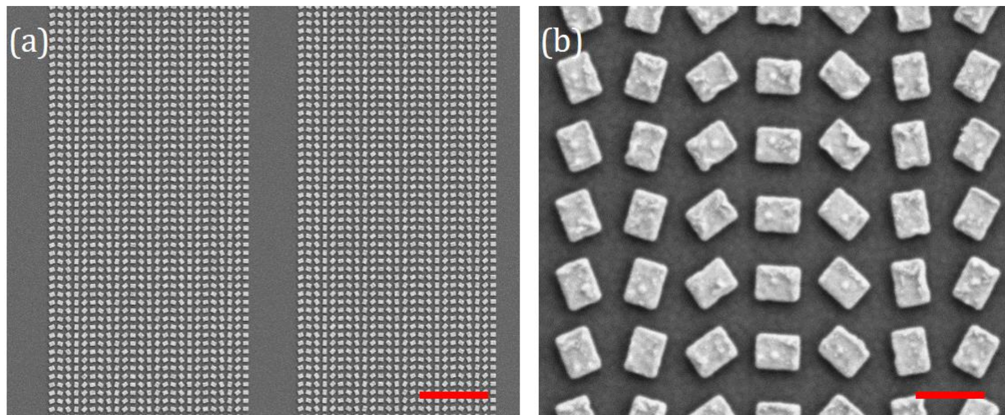

Figure 5.4: SEM images of a metasurface patterned by EBL. Scale bars in (a) and (b) are 5 μm and 600 nm, respectively.

6) Titanium cover

   In each device, a rectangular titanium layer with a square window on the central area, serving the purpose of preventing leakage of the incident laser beam into the substrate around the metasurface area. This titanium cover is fabricated with the standard process of photolithography, e-beam evaporation, and lift-off.

7) Top windows for wire bonds

   Next, a large area of the silica layer on top of each Au contact pad is removed, to allow for wire bonding in the packaging step. Similar to step 2, photolithography and BOE etch are used to pattern these windows. A photograph of a completed wafer after this step is shown in Figure 5.5.

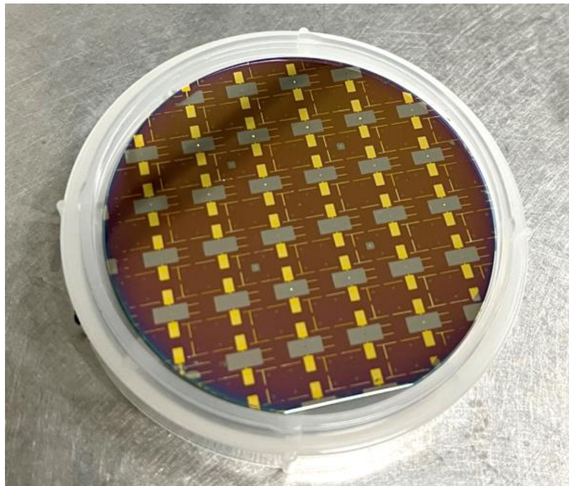

Figure 5.5: Photograph of a completed wafer after the 7 fabrication steps before dicing. The wafer contains 26 complete spaces for devices of 7.5-mm lateral dimensions to be diced into 26 chips.

8) Wafer dicing

   Before dicing the wafer, a protective photoresist layer is applied by spin coating and baking. Specifically, two layers of S1818 G2 resist spin-coated at 3000 rpm have been found to be sufficient for this purpose. The wafer is then diced into 7.5-mm by 7.5-mm chips (this size is determined by the photomask designs, as discussed in section 2.4) using the Disco DAD 3220 dicing tool with a diamond hub blade. The protective photoresist is then removed with acetone.

9) Evaluation

   Due to the labor-heavy and time-consuming FIB operation in the next step, sample selection is strongly recommended to screen out samples showing structural damage and those with short or open circuits. Using an optical microscope, the visual appearance of each chip can be evaluated, with special attention paid to the integrity of the Ti cover and the metasurfaces. Next, the DC resistance of each device that passed the visual inspection is measured with a home-built probe station. An optimal value of the resistance is between 1000 and 20,000 ohms, although devices with smaller resistance of a few 100 ohms have also produced good results in the past.

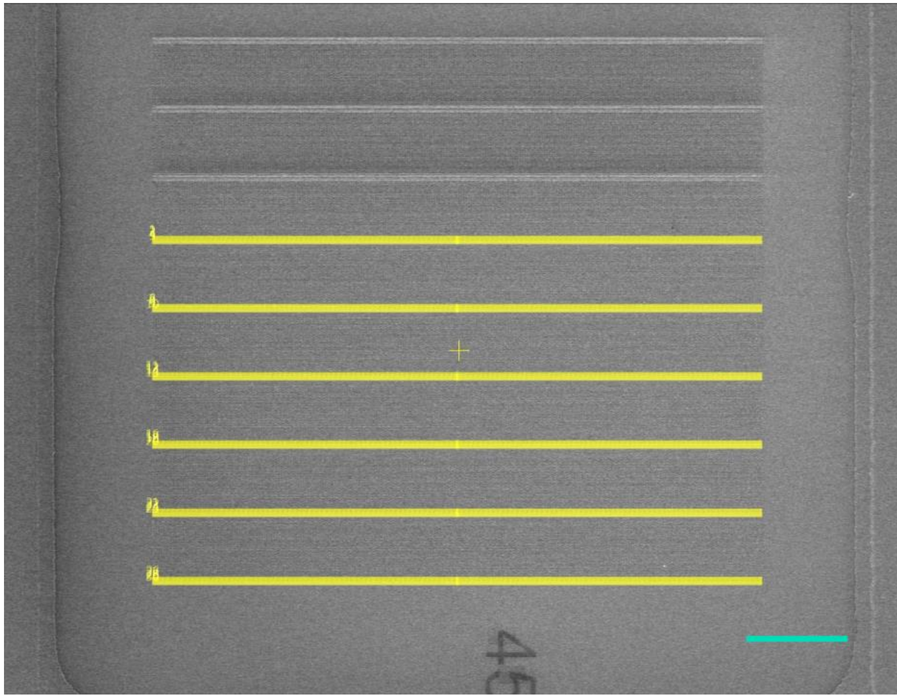

Figure 5.6: FIB patterning window. The greyscale image is a snapshot by the ion beam, and the yellow contours are the slit patterns in progress. Near the top of the image, three groups of slits are clearly visible. The scale bar is 20 um.

10) Slit patterning

FIB with a gallium source is then performed to pattern the slits through the topmost silica and gold layers. This process is quite straightforward, because one can visually align the pattern to be milled on the existing structures with the help of SEM images and ion-beam scans (see Figure 5.6).

11) Ball bonding and packaging

The chips with complete nanostructures are finally packaged for the subsequent measurements. Each chip is attached by double-sided copper tape to a metal block, and two ceramic pads with gold coatings on their top surfaces are then

similarly attached on both sides of the chip. Next, ball bonding is conducted to connect the contact pads to the chip electrodes with gold wires.

5.2.3 Recipes for Important Steps

The fabrication processes described in the previous section employ the detailed recipes below besides the standard operation procedure of the relevant machines. These recipes are specifically summarized for the metasurface devices discussed in this thesis, according to past successes and failures.

1. Photolithography.

   All the photolithography steps involve positive-tone exposure, i.e. the photoresist is removed by the developer wherever it is exposed to UV light. These are the detailed instructions.

   At room temperature, spin coat Microposit S1813 G2 photoresist onto the wafer using the Headway Research PWM32. The spinning recipe is as follows:

   500 rpm speed, 500 rpm/s ramp, 5 sec;

   4000 rpm speed, 1000 rpm/s ramp, 45 sec;

   500 rpm speed, 1000 rpm/s ramp, 5 sec;

   0 speed, 500 rpm/s ramp, 0 sec.

   This results in 1.2 um thickness.

   Then, immediately bake the wafer on a hot plate at 115 degrees for 1 min to solidify the photoresist.

   After cool-down, align and expose UV light at 405 nm line with Karl Suss MA/BA-6 mask aligner. The exposure recipe is:

   Dose: 150 mJ/cm^2, channel 1;

   Alignment: contact mode, 50 um gap;

Exposure sequence: low-vacuum; pre-vacuum 0 sec, vacuum sequency 4 sec, vacuum purge 4 sec;

To develop the photoresist, immerse the wafer in MF319 and shake constantly for 30~45 sec, and then rinse twice with DI water. Blow dry the surface with a nitrogen gun.

In certain cases, a post bake is required in order to remove the water molecules absorbed into the material at the top surface.

2. Metal deposition.

   All the metals on the wafer involved in this process are deposited by the CHA Industries Solution machine, with a technique of physical-vapor deposition using an electron beam as the heating source. The adhesion layer of titanium is evaporated at a rate of 0.5 A/s, and gold of 99.999% purity is evaporated at 1.0 A/s. A vacuum level of at least 10^{-6} is necessary to guarantee high quality films. AFM measurements reveals that such a recipe can produce thin films of ~2.5 nm RMS roughness.

3. EBL.

   Spin coat PMMA 950 A4 (solution in anisole) on the wafer following the recipe below:

   500 rpm speed, 500 rpm/s ramp, 5 sec;

   4000 rpm speed, 1000 rpm/s ramp, 45 sec;

   500 rpm speed, 1000 rpm/s ramp, 5 sec;

   0 speed, 500 rpm/s ramp, 0 sec.

This will form a layer of PMMA with about 200 nm thickness.

Then, immediately bake the wafer on a hot plate at 180 degrees for 15 min or more to evaporate the solvent.

After cool down, expose the pattern with high-energy electron beam. For example, the metasurfaces introduced in this thesis can be patterned with Elionix HS-50 with the following recipe:

Beam current: 2 nA;

Dose: 300 uC/cm^2 at 50 keV;

Writing field: 250 um;

Resolution: 4 nm.

To develop the e-beam resist, immerse the wafer in MIBK/IPA 1:3 and shake constantly for 60 sec, and then rinse with IPA. Blow dry the surface with a nitrogen gun.

4. PECVD.

STS PECVD machine by Surface Technology Systems is used in this processing step. According to the calibrated deposition rate, the built-in recipe of high-frequency silicon oxides (HFSIO) for 1:08 min/sec results in a 60-nm-thick SiO_2 conformal layer, and the refractive index is about 1.47. Similarly, to achieve a 120-nm SiO_2 film, an HFSIO recipe of 2:16 min/sec is needed. PECVD requires a substrate temperature of 200 degrees, so any trapped water inside the deposited layers is hazardous to the quality of samples.

5. FIB.

FEI Helios 660 tool is used to perform FIB. The recipe for making 5 200-nm-wide slits with a period of 400 nm is listed here:

Ion-beam: gallium source, 30 kV, 230 pA;

Target material: Si;

Magnification: 800;

Dimension of each slit pattern (setting in the control software): 150 um by 150 nm by 160 nm (length/width/depth).

The slits shall be patterned horizontally in order to minimize the effect of machine drift and ensure a smooth profile of the slits. As a compensation technique, it is recommended to start with patterning 3 groups of slits and use SEM to observe the machine drift. In addition, the EBL designs mentioned in section 5.2.2 should include supplementary features to allow for the alignment and calibration of both electron and ion beams in this step.

6. BOE etch

The BOE solution of 1:6 dilution at room temperature etches oxides at a rate of 28 A/s according to the manufacturer's specification. Therefore, it's recommended to etch at least 45 sec for 60-nm and 1:30 min/sec for 120-nm SiO_2 films.

Usage of BOE asks for extreme caution and a specific training for the personnel involved. Furthermore, a specialized set of wafer handling tools (tweezers, colander, and dishes) is needed because hydrofluoric acid may corrode regular tools. These are the detailed instructions.

Immerse the wafer in BOE for a timed period with the container constantly shaken, and quickly transfer to DI water to rinse it twice. Blow dry the wafer with a nitrogen gun.

7. Photoresist removal/lift off

   There are two ways to remove photoresist:

   a) Put the sample in sufficient amount of acetone for about 3 minutes, or until the colored photoresist fully dissolves in the solvent;

      Transfer to IPA and rinse;

      Quickly use N_2 to blow dry.

   b) Use the mask aligner to perform flood exposure with 405-nm UV light for a sufficient amount of time;

      Soak the sample in developer MF319 for about 2 minutes, or until the colored photoresist fully goes away;

      Transfer to DI water and rinse twice;

      Quickly use N_2 to blow dry;

      Bake on a hot plate at 110 degrees for 2 minutes to remove moisture.

   To do lift off:

   Immerse the sample in a covered and heated acetone bath for 15~30 minutes, or until the metal film starts to peel off by itself;

   Shake the container and use an acetone wash bottle to help lift off the film;

   If the film is not completely detached from the substrate, sonicate in the same container for 5~15 seconds;

Transfer to another acetone container and rinse, in order to further clean up the metal debris;

Transfer to a container filled with IPA and rinse;

Take the wafer out and quickly use a nitrogen gun to dry.

5.2.4 Photomasks Preparation

In this new fabrication process, there are six photolithography steps that require new and distinctive photomasks for positive-tone exposure. The six mask designs, drawn with AutoCAD, are presented in Figure 5.7. Upon designing, it was taken into consideration that larger wafer sizes might be used in the future, so an area of 10 by 10 centimeters was covered with an array of patterns with a corresponding photomask size of 5 inches.

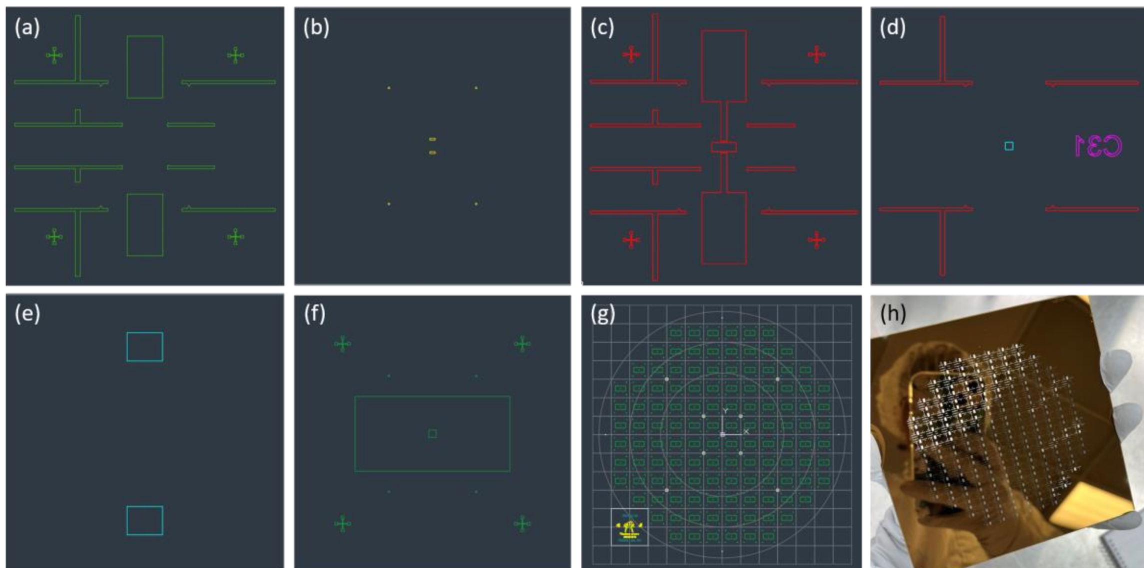

Figure 5.7: Photomask designs. (a) – (f) Photolithography patterns at the chip scale for steps (1) – (7) in section 5.2.2. (g) Full-size photomask design. (h) A completed photomask.

The photomasks used are made of soda-lime glass and chrome, with a lateral dimension of 5 by 5 inches. A photoresist layer of type AZ1518 is already coated and baked on the chrome side by the supplier. The maskless writer Heidelberg MLA-150 is used to expose the photoresist, and then developed with AZ-400K. A wavelength of 405 nm is chosen, and the laser write head produces a resolution of 1 um. Then,

chrome is etched with a liquid etchant, and the photoresist is removed by Remover 1165.

5.3 Probe Station

5.2.1 Introduction to the Design

In our previous sample fabrication process, the electrical evaluation was conducted after sample packaging and ball bonding, by directly using a multimeter to touch the bonding pads and measure the resistance. However, such method becomes outdated after the whole-wafer upgrade. An immediate drawback is the uncertainty of manual operations when using hands to control the probes. This may damage the device contact pads and result in unreliable readings. At the same time, with the large increase in number of fabricated samples enabled by the process upgrade, it is worth developing a method to evaluate and select the samples before packaging. As a result, I constructed a probe station for 2-probe DC measurements as a peripheral tool.

The self-built probe station has 4 major components, as shown in Figure 5.8. The default probes are made from sections of an iron wire and are clamped by the adjustable arms. They can be changed to other materials according to the user preference. The internal resistance of the system is about 1 ohm. There are 4 moving stages: microscope, sample, and the two probe arms. All the stages allow for 3-dimensional movement (the microscope z-axis controls the focus).

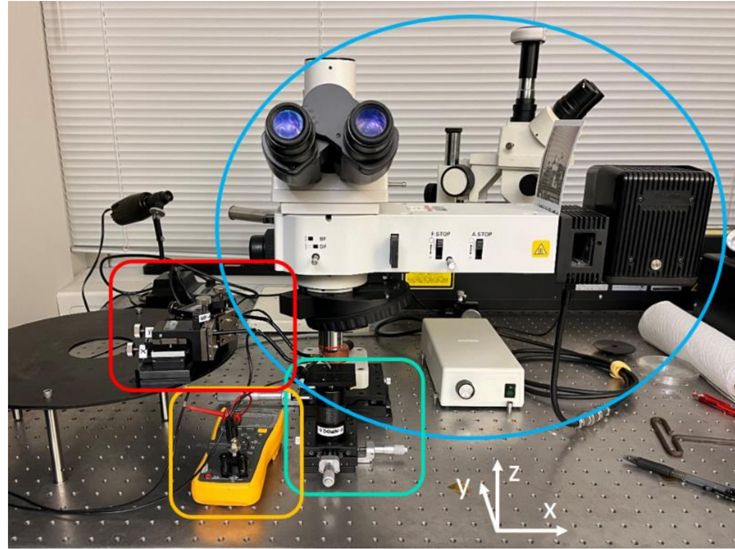

Figure 5.8: Front view of the probe station. Red rectangle: probes and stages, used to direct the metallic needles to make contact with the sample. Blue rectangle: microscope and lamp, used to observe the sample and probe tips in order to navigate and focus. Yellow rectangle: multimeter, used to measure the DC resistance. Green rectangle: sample stage for holding and moving the sample in 3 dimensions.

5.2.2 Standard Operation Procedure

It is suggested to use the left hand to control the microscope x- and z-axes, the probes movement, and the multimeter, while the right hand controls the sample stage and the microscope y-axis. The standard operation procedure is as follows:

1) Pre-operation check:

   Upon arriving, check that the working area is clean, and that there are no additional wires or random stuff that block the setup operation or stage movement;

   Ground yourself by touching the table with both hands;

   Check that all the stages have sufficient freedom of movement;

   Check that the microscope can move down but doesn't touch the probes.

2) Turning on the system:

   Turn on the lamp, and tune its brightness to about ½;

   Turn on the multimeter to a reasonable measurement range (e.g., 1k Ohms or 10k Ohms).

3) Focusing on the tips:

   Move the microscope around to find one of the probes tip and focus on it (use either a light-colored paper or a sample as the background);

   Move the microscope and the other probe so that they both appear in the field-of-view;

   Adjust the height of the other probe so its tip also comes to the focal plane (see Figure 5.9).

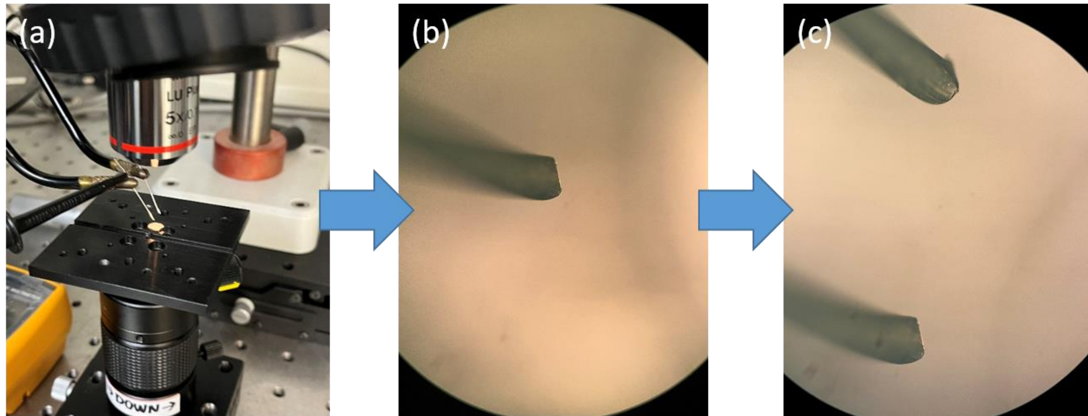

Figure 5.9: Workflow of focusing on the tips.

4) Loading the sample:

   Lower the stage and put the sample on it (making sure that the sample doesn't slide around easily), roughly around the field of view of the microscope;

   Slowly raise the sample stage and stop at around 1mm away from the probes. Move the sample stage so that the probes are close to the contact pads;

   Lower the microscope to adjust the focus until you see the patterns on the sample clearly.

5) Making contact:

   Move the sample stage so that you are looking at one contact pad, and then fine adjust to let one probe lie on the target location;

   Use the z-axis knob to drop the probe and make contact (the tip slides a little on the surface and then it comes into focus);

   Move the microscope to the other contact pad, and move the other probe to the target location;

   Drop the probe to make contact.

6) Measurement:

   Read the signal on the multimeter after it becomes stable. If there is no signal, try changing its measurement range.

7) Finalization:

   After obtaining the data, raise the probes while observing with the microscope, so that both probes lose contact from the surface;

   Turn the z-knobs of the probes upward by 2 turns (this raises about 1 mm);

   Lower the sample stage to a reasonable distance.

8) Additional measurements:

   If testing another sample, move the sample or the sample stage and repeat steps (4) – (7).

9) Wrap-up:

   Turn down and turn off the lamp;

   Turn off the multimeter.

CONCLUSION

Throughout this Dissertation work, my research has focused on the development of a new type of optical detectors that are sensitive to the angle of the incident light, and on the special imaging capabilities of pixel arrays of these devices, including spatial filtering and quantitative phase imaging. The designs that I developed are based on plasmonic metasurfaces patterned on top of a multi-layer structure, which can be directly integrated on standard image sensors. Detailed design parameters were determined by Ansys FDTD simulations to optimize performance in terms of target detection angles, peak transmission, and transmission slope. Experimental samples were fabricated via a variety of cleanroom techniques, and the process was recently upgraded to a wafer-scale level, resulting in a major improvement in yield, consistency, and accuracy. These samples were then characterized with a simple measurement setup to obtain their angular response. I also introduced a rigorous theoretical model based on Fourier optics to study the image processing capabilities of these devices, under both coherent and spatially incoherent illumination. Our proposed devices, compared to existing solutions employing traditional optical components, have many advantages including extreme size miniaturization, ease of alignment, pixel-by-pixel operation flexibility, and simpler fabrication (2-dimensional patterning).

Specifically, in order to enable the function of edge-enhanced imaging, I designed and developed two devices with symmetric angular response around normal incidence. In the spatial frequency domain, an array of such devices with small

nonzero angle of peak detection is capable of performing band-pass filtering, as needed for edge enhancement with coherent light. Under incoherent illumination, edge detection could be obtained by combining directional devices with standard pixels on a same array, and taking the difference between their respective images. To further enable phase-contrast imaging, I discussed the approach of building metasurfaces with asymmetric response near normal incidence. When inserted into a standard microscope imaging system, an array of such sensors is capable of resolving a phase contrast as low as 2 mrad. With the help of a special pixel arrangement and reconstruction algorithm, I demonstrated the ability of quantitative phase reconstruction in a single shot. Finally, for other applications where larger detection angles are required (such as compound-eye vision with ultrawide field of view), I developed true gradient-metasurface devices based on aperiodic nanostripe arrays that not only feature higher peak transmission efficiency but also open up new design opportunities.

Although all the devices reported in this Dissertation operate at a near-infrared wavelength of 1550 nm, the same design method can be extended to the visible range by using all-dielectric materials. For instance, dielectric photonic-crystal slabs can be used as waveguides as well as reflection mirrors (similar to the metal films in the current devices), and dielectric nanoparticles placed on top of the slab can then be used to promote directional photodetection. Based on the same design platform, metasurfaces with more complex nanoparticle geometries and arrangements may be designed to enable broadband achromatic operation. This goal could be achieved

either by introducing a sufficiently large number of design parameters to engineer the angular-response dispersion, or by taking advantage of machine learning. Similarly, 2-dimensional metasurfaces could be developed to achieve polarization-independent response, by separating the two orthogonal polarization components of the incident light and guiding them into SPPs travelling along two different directions. In addition, the first layer of convolutional neural networks for image recognition tasks can be implemented with our devices in the optical domain, as a way to improve speed and reduce power consumption. For this purpose, multiple channels of optical spatial filtering or phase contrast imaging can be performed simultaneously by constructing partitioned sensor arrays with directional photodetectors of different orientations. In recent collaborative work, we have applied this idea to the classification of lung cancer cells of 7 different subtypes, and our results show comparable performance to a reference fully digital network with 20× reduction in number of operations. Finally, the improved fabrication process that I completed in the course of this work could be extended further to realize batch production of our proposed devices directly on commercial image sensor arrays.

APPENDIX A

CTF Calculation

Here we extend the model presented in Chapter 2 Section 5 for the CTF of our plasmonic directional image sensors, to allow for a finite y component k_y of the incident wavevector **k** and for the presence of slits on both sides of the NP array. We consider the structure shown schematically in Figure A.1, which consists of N rectangular NPs oriented along the y direction, arranged periodically at positions x_l (l = 1, 2, 3, …, N) with period $\Lambda = x_l - x_{l-1}$, and with two slits located symmetrically at $x_{sl-} = x_1 - \delta$ and $x_{sl+} = x_N + \delta$. The device is illuminated with a harmonic plane wave of in-plane wavevector $\mathbf{k} = \hat{\mathbf{x}}k_x + \hat{\mathbf{y}}k_y$, so that the incident field distribution on the NP array is $E_{in,\mathbf{k}}(\mathbf{r}) = E_{in}(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{r}}$. The incident light is scattered by each NP into a superposition of plane waves of different wavevectors, subject to the constraint that k_y is conserved due to the relatively large NP length along the y direction. These scattered waves can then excite SPPs on the underlying metal film if the requirements of energy and momentum conservation are satisfied.

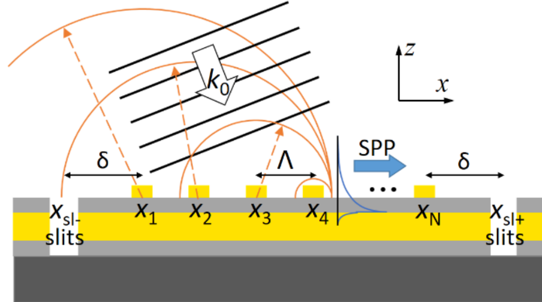

Figure A.1: Schematic illustration of the physical model used to evaluate the CTF.

With this description, the SPP field at the slit positions can be expressed as

$$\mathrm{E}_{\mathrm{SPP},\mathbf{k}}(\mathbf{r}_{\mathrm{sl}\pm}) = \sum_{l=1}^{\mathrm{N}} \int_{-\mathrm{w}/2}^{\mathrm{w}/2} \frac{\mathrm{d}y_l}{\mathrm{w}} \mathrm{E}_{\mathrm{in},\mathbf{k}}(\mathbf{r}_l) \int \mathrm{d}\tilde{\mathrm{k}}\, \eta_{\mathrm{SPP}\pm}\big(\mathbf{k} \pm \hat{\mathbf{x}}\tilde{\mathrm{k}}\big) \mathrm{e}^{\mathrm{i}\left(\mathbf{k}\pm\hat{\mathbf{x}}\tilde{\mathrm{k}}\pm\mathrm{i}\hat{\mathbf{x}}\gamma\right)\cdot\left(\mathbf{r}_{\mathrm{sl}\pm}-\mathbf{r}_l\right)}, \quad \text{(A.1)}$$

where each term in the sum over the index l is the contribution from the NP located at position $\mathbf{r}_l$, averaged over the NP length w along the y direction. In the integral over $\tilde{\mathrm{k}}$, $\eta_{\mathrm{SPP}\pm}\big(\mathbf{k} \pm \hat{\mathbf{x}}\tilde{\mathrm{k}}\big)$ is the probability amplitude that the light scattered by each NP with in-plane wavevector $\mathbf{k} \pm \hat{\mathbf{x}}\tilde{\mathrm{k}}$ excites a SPP that propagates towards the slit at $\mathrm{x}_{\mathrm{sl}\pm}$. Its $\mathbf{k}$ dependence is determined by the phase matching condition including the SPP finite propagation length $\mathrm{L}_{\mathrm{SPP}}$, so that $\left|\eta_{\mathrm{SPP}\pm}\big(\mathbf{k} \pm \hat{\mathbf{x}}\tilde{\mathrm{k}}\big)\right|^2$ is a Lorentzian function of $\mathrm{k}_\mathrm{x} \pm \tilde{\mathrm{k}}$ centered at $\pm\sqrt{\mathrm{k}_{\mathrm{SPP}}^2 - \mathrm{k}_\mathrm{y}^2}$ (where $\mathrm{k}_{\mathrm{SPP}}$ is the SPP wanumber at the illumination wavelength λ_0) with full width at half maximum (FWHM)) $1/\mathrm{L}_{\mathrm{SPP}}$. The exponential factor in Equation (A.1) accounts for the phase shift and attenuation experienced by the excited SPPs as they travel from their points of origin $\mathbf{r}_l$ to the slits, with attenuation factor $\gamma = 1/(2\mathrm{L}_{\mathrm{SPP}})$. For simplicity, in the attenuation term the distance traveled by each SPP is approximated as $\hat{\mathbf{x}} \cdot (\mathbf{r}_{\mathrm{sl}\pm} - \mathbf{r}_l)$, rather than using the exact value $|\mathbf{r}_{\mathrm{sl}\pm} - \mathbf{r}_l|$. This approximation is valid as long as $\mathrm{k}_\mathrm{y} \ll \mathrm{k}_{\mathrm{SPP}}$ (so that the excited SPPs travel at a small angle with respect to the x axis), which is always the case in the low-numerical-aperture imaging systems considered in this work.

Using $\mathrm{E}_{\mathrm{in},\mathbf{k}}(\mathbf{r}_l) = \mathrm{E}_{\mathrm{in}}(\mathbf{k})\mathrm{e}^{\mathrm{i}\mathbf{k}\cdot\mathbf{r}_l}$ in the right-hand side of Equation (A.1), we find that the resulting expression does not depend on y_l, and therefore the averaging over the length of each NP along the y direction can be dropped. The equation can be further simplified using $\mathrm{x}_{\mathrm{sl}+} - \mathrm{x}_l = (\mathrm{N} - l)\Lambda + \delta$ and $\mathrm{x}_{\mathrm{sl}-} - \mathrm{x}_l = -(l - 1)\Lambda - \delta$. With these

substitutions, we finally arrive at the expression of Equations (3) and (4) of the main text for the CTF $t_{\pm}(\mathbf{k}) = E_{SPP,\mathbf{k}}(\mathbf{r}_{sl\pm})/E_{in,\mathbf{k}}(\mathbf{r}_{sl\pm})$, i.e.,

$$t_{\pm}(\mathbf{k}) = \int d\tilde{k}\, \eta_{SPP\pm}(\mathbf{k} \pm \hat{\mathbf{x}}\tilde{k}) e^{i(\tilde{k}+i\gamma)\delta} f(\tilde{k}), \tag{A.2}$$

where

$$f(\tilde{k}) = \left[1 - e^{i(\tilde{k}+i\gamma)N\Lambda}\right] / \left[1 - e^{i(\tilde{k}+i\gamma)\Lambda}\right]. \tag{A.3}$$

Based on the definition above for the probability amplitude $\eta_{SPP\pm}$, we can also argue that $\eta_{SPP-}(\mathbf{k} - \hat{\mathbf{x}}\tilde{k}) = \eta_{SPP+}(-\mathbf{k} + \hat{\mathbf{x}}\tilde{k})$. It then follows from Equation (A.2) that

$$t_{-}(\mathbf{k}) = t_{+}(-\mathbf{k}), \tag{A.4}$$

which is consistent with the symmetric device geometry under study.

To calculate theoretical CTF shown in Figure 2.5, I have developed a MATLAB code and it is attached below:

```
%This code simulates t(k) with the theoretical 2D model of the metasurface

%input: 'phase_Ey', is the nano-particle's right-SPP-scattering phase
%   vs incidence angle.
% size: (1,181), 180 degrees range
% example: phase_Ey = zeros(1,181);
%Notice: k_prime space and k-space should have the same sampling rate!
%Important: 1.2 factor should be found
%all dimensions are in meters.


%% input arguments:
Lspp = 80e-6;
delta = 1.485e-6;               %geometric parameter
Lambda = 1.485e-6;              %grating period, in microns
N = 15;                         %grating number
alpha = 0;                      %phase accumulation of SPP passing each NP

%the following arguments should correspond to OTF_calculation code:
lambda = 1550e-9;
data_size = 401;
interp_num = 4;
```

```
extension = 1.3;                    %calculating how big range of k relative to k0
M = interp_num*extension*(data_size-1)+1;

qmax = 5.7762e5;

k0 = 2*pi/lambda;
kspp = 1.0438*k0;
gamma = 1/2/Lspp;                   %gamma is loss coefficient
%Lorentzian function with FWHM = 1/Lspp
FWHM = 1/Lspp;

k = linspace(-extension*k0,extension*k0,M);
%choose a range of k_prime of at least 3.5 times the size of (-k0,k0):
M_prime = (M-1)/extension*3.5+1;        %number of k' to be integrated
k_prime = linspace(-3.5*k0,3.5*k0,M_prime);
dk_prime = k(2)-k(1);


%%discrete integral:
tk_plus = zeros(1,M);
ETA_index = zeros(1,M);
for i=1:M
   eta = sqrt(2/pi*FWHM./(4*(k(i)+k_prime-kspp).^2+FWHM^2));
   fk = (1-exp(1i*N*((k_prime+1i*gamma)*Lambda+alpha)))./(1-
exp(1i*((k_prime+1i*gamma)*Lambda+alpha)));
   tk_plus(i) = dk_prime*sum(eta.*exp(1i*(k_prime+1i*gamma)*delta).*fk);
end

tk_plus = tk_plus/max(tk_plus);

figure;
plot(k,abs(tk_plus));
hold on;
grid on;
plot(k,angle(tk_plus));

figure;
T_calc = (abs(tk_plus)).^2;
plot(k,T_calc);
grid on;
title('calculated T');
```

To further substantiate the predictions of the analytical model, including the linear k_x-dependence of $\phi_{\pm}(\mathbf{k})$, we have also modeled the CTF $t_{+}(k_x,k_y=0)$ via FDTD

simulations. In these calculations, we consider a 2D structure consisting of the full 15-NP array of our experimental device A but without any slits, with PMLs on all 4 boundaries. The incident light is a p-polarized Gaussian beam of variable angle of incidence θ, focused on the device surface with spot size of 20.8 μm (close to the full size of the NP array) and small convergence angle 2.7°. With this setup, we have computed the surface-normal component of the electric field E_z of device A (i.e., the dominant SPP field component) at a distance δ = 41.6 μm from the center of the rightmost NP (a distance far enough where the incident beam is negligible) as a function of $k_x = (2\pi/\lambda_0)\sin\theta$. The calculated magnitude of E_z versus k_x is plotted in Figure A.2(a) and shows good agreement with the predictions of the analytical model for $|t_+(k_x,k_y=0)|$ presented in Figure 2.5(a) of the main text. In Figure A.2(b) we plot the simulated phase of E_z versus k_x, referenced to the phase of the incident electric field at the the location of the slit at x_{sl+}. The linear k_x-dependence predicted by the analytical model around $k_x = 0$ is also well reproduced by these simulation results. The value of the slope parameter $\alpha = -d\phi_+(\mathbf{k})/dk_x$ inferred from Figure A.2(b) is 12.2 μm, reasonably close to the analytical estimate of 9.8 μm discussed in the main text. The difference between these two values is ascribed to the smaller SPP propagation length inferred from the experimental data and used in this analytical estimate compared to the FDTD simulations, likely due to SPP scattering by surface roughness in the actual samples.

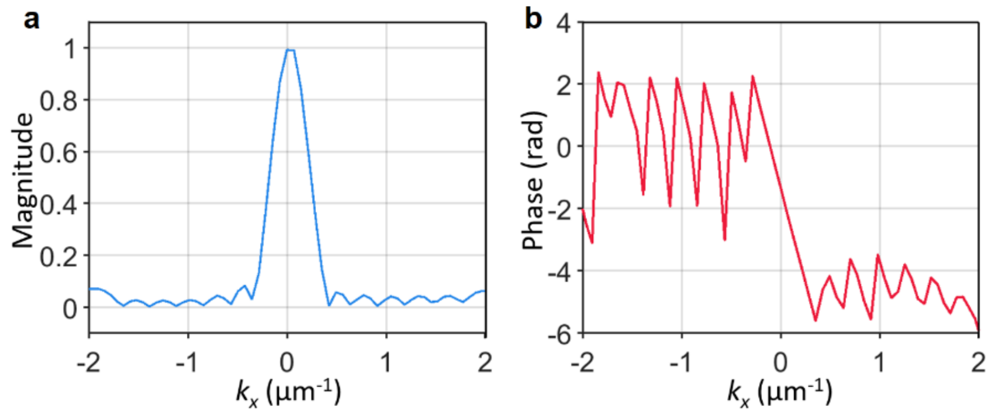

Figure A.2: FDTD simulation results. Normalized magnitude (a) and phase (b) of the CTF $t_+(k_x,k_y=0)$ versus k_x, computed with the parameter values of sample A.

Similar derivation is also applied to Chapter 3 Section 4 to obtain the phase slope $\alpha = 8$ μm.

APPENDIX B

OTF Calculation

B.1 Derivation of the Filter OTF

Under spatially incoherent illumination, the correlation function of the incident light $\langle E_{in}^*(\mathbf{r}-\frac{\delta r}{2})E_{in}(\mathbf{r}+\frac{\delta r}{2})\rangle$ (where the brackets $\langle\ldots\rangle$ indicate an ensemble average) is a monotonically decreasing function of distance δr between points on the illuminated surface, with maximum value at $\delta r = 0$ determined by the input intensity $I_{in}(\mathbf{r})$. Here we consider the simple Gaussian model

$$\langle E_{in}^*(\mathbf{r}-\tfrac{\delta r}{2})E_{in}(\mathbf{r}+\tfrac{\delta r}{2})\rangle \propto I_{in}(\mathbf{r})\exp(-\delta r^2/4\Delta^2)/\Delta, \tag{B.1}$$

which can be readily converted to the spatial-frequency-domain expression

$$\langle E_{in}^*(\mathbf{k})E_{in}(\mathbf{k}')\rangle \propto I_{in}(\mathbf{k}'-\mathbf{k})\exp(-\Delta^2|\mathbf{k}'+\mathbf{k}|^2/4)\Delta, \tag{B.2}$$

where again $E_{in}(\mathbf{k})$ and $I_{in}(\mathbf{k})$ are the Fourier transforms of the incident field distribution $E_{in}(\mathbf{r})$ and optical intensity $I_{in}(\mathbf{r})$, respectively. The transverse coherence length Δ introduced in these equations is discussed in section B.2 below.

In order to evaluate the OTF $T_{filter}(\mathbf{q}) = I_{out}(\mathbf{q})/I_{in}(\mathbf{q})$ of an optical spatial filter, we begin by expressing the output intensity distribution $I_{out}(\mathbf{r}) \propto \langle E_{out}^*(\mathbf{r})E_{out}(\mathbf{r})\rangle$ in terms of the Fourier transform of the output field $E_{out}(\mathbf{k}) = t(\mathbf{k})E_{in}(\mathbf{k})$. With this approach, we obtain

$$I_{out}(\mathbf{r}) \propto \iint d\mathbf{k}'d\mathbf{k}\langle E_{out}^*(\mathbf{k})E_{out}(\mathbf{k}')\rangle\exp(i(\mathbf{k}'-\mathbf{k})\cdot\mathbf{r}) =$$

$$= \iint d\mathbf{q}d\mathbf{k}\; t^*(\mathbf{k})t(\mathbf{k}+\mathbf{q})\langle E_{in}^*(\mathbf{k})E_{in}(\mathbf{k}+\mathbf{q})\rangle\exp(i\mathbf{q}\cdot\mathbf{r}), \tag{B.3}$$

where we have made the change of variable $\mathbf{q} = \mathbf{k}'-\mathbf{k}$.

By combining Equations (B.2) and (B.3) and using $\mathrm{I_{out}}(\mathbf{r}) = \int \mathrm{d}\mathbf{q}\, \mathrm{I_{out}}(\mathbf{q})\exp(\mathrm{i}\mathbf{q}\cdot\mathbf{r}) = \int \mathrm{d}\mathbf{q}\, \mathrm{T_{filter}}(\mathbf{q})\mathrm{I_{in}}(\mathbf{q})\exp(\mathrm{i}\mathbf{q}\cdot\mathbf{r})$, we finally find

$$\mathrm{T_{filter}}(\mathbf{q}) \propto \int \mathrm{d}\mathbf{k}\, \mathrm{t}^*(\mathbf{k})\mathrm{t}(\mathbf{k}+\mathbf{q})\exp(-\Delta^2|2\mathbf{k}+\mathbf{q}|^2/4). \tag{B.4}$$

This expression is the same as Equation (2.6) of the main text for the OTF $\mathrm{T_{filter}}(\mathbf{q}) = \mathrm{T_{\pm}}(\mathbf{q})$ in terms of the CTF $\mathrm{t}(\mathbf{k}) = \mathrm{t_{\pm}}(\mathbf{k})$.

With this equation and using the change of variable $\tilde{\mathbf{k}} = \mathbf{k} + \mathbf{q}$, one can readily show that $\mathrm{T}^*_{\pm}(\mathbf{q}) = \mathrm{T_{\pm}}(-\mathbf{q})$, which is a general result. Furthermore, for the symmetric devices under study where $\mathrm{t_-}(\mathbf{k}) = \mathrm{t_+}(-\mathbf{k})$, the same equation with the change of variable $\tilde{\mathbf{k}} = -\mathbf{k} - \mathbf{q}$ yields

$$\mathrm{T_-}(\mathbf{q}) = \mathrm{T}^*_{+}(\mathbf{q}), \tag{B.5}$$

which is used in the derivation of Equation (2.8) of the main text. By combining these two conditions, we also find that in symmetric devices $\mathrm{T_-}(\mathbf{q}) = \mathrm{T_+}(-\mathbf{q})$.

B.2 Estimation of the Transverse Coherence Length

The transverse coherence length introduced in Equation (B.1) is determined following considerations based on the van Cittert – Zernike theorem [S1]. In this formulation, if the sensor array is illuminated with light from a fully incoherent object of angular diameter θ_s, the radius of coherence r_c on the sensor plane is equal to $1.22\lambda_0/\theta_s$ as a result of the gain in coherence upon free-space propagation. For an object at a representative distance of 10 times its lateral size, i.e., with θ_s = 2arctan(0.05) ≈ 0.1 rad, r_c is equal to 18.9 μm at λ_0 = 1.55 μm. In the Gaussian model of Equation (B.1), r_c can be approximated as the half-width-at-1/e^2-maximum $2\sqrt{2}\Delta$ of the incident-light correlation function, so that a radius of coherence r_c = 18.9 μm corresponds to a value of 6.7 μm for Δ. In practice, this value should be regarded as a lower bound because it neglects the finite coherence length of the light scattered by the object. It should also be noted that, in the plasmonic devices under study, the radius of coherence should be larger than the half-size of the NP array $(N-1)\Lambda/2$ to allow for constructive interference among the SPPs scattered by all the NPs. In our experimental devices, $(N-1)\Lambda/2 \approx 11$ μm so that this condition is fully satisfied for r_c = 18.9 μm.

B.3 Calculation of the Differential OTF

The pixel array of Figure 2.5(a) of the main text is described by the differential OTF $\Delta T_{tot}(\mathbf{q}) = \Delta T_{meas}(\mathbf{q}) T_{lens}(\mathbf{q}) = \left[T_{meas}^{ref}(\mathbf{q}) - T_{meas}^{plas}(\mathbf{q})\right] T_{lens}(\mathbf{q})$, where $T_{meas}^{ref}(\mathbf{q})$ and $T_{meas}^{plas}(\mathbf{q})$ are the measurement OTFs of the reference (uncoated) and plasmonic pixels, respectively, and $T_{lens}(\mathbf{q})$ is the lens transfer function of Equation (2.9). $T_{meas}^{plas}(\mathbf{q})$ can be computed using Equation (2.8) of the main text. In order to derive a similar expression for $T_{meas}^{ref}(\mathbf{q})$, we begin by noting that each reference pixel simply integrates the optical intensity $I_{in}(\mathbf{r})$ incident on its illumination window. The responsivity of these devices does not feature any significant angular dependence, other than the decrease in transmission through the illuminated surface at large angles of incidence (well beyond the numerical aperture of the imaging systems considered in this work) predicted by the Fresnel formulas. Therefore, the photocurrent measured by the $\boldsymbol{n}^{th}$ reference pixel is proportional to

$$I_{meas}^{ref}(\boldsymbol{n}) = \int_{-w_x/2}^{w_x/2} \frac{dx}{w_x} \int_{-w_y/2}^{w_y/2} \frac{dy}{w_y} I_{in}(x_c^{\boldsymbol{n}} + x, y_c^{\boldsymbol{n}} + y), \tag{B.6}$$

where w_x and w_y denote the size of the illumination window along the x and y directions, respectively, and $\mathbf{r}_c^{\boldsymbol{n}}$ is the pixel center position. Next, we express $I_{in}(\mathbf{r})$ in terms of its Fourier transform $I_{in}(\mathbf{q})$, evaluate the two integrals, and finally extract the Fourier transform of $I_{meas}^{ref}(\boldsymbol{n}) = \int d\mathbf{q} I_{meas}^{ref}(\mathbf{q}) e^{i\mathbf{q}\cdot\mathbf{r}_c^{\boldsymbol{n}}}$. With this procedure we find that the OTF $T_{meas}^{ref}(\mathbf{q}) = I_{meas}^{ref}(\mathbf{q}) / I_{in}(\mathbf{q})$ is given by

$$T_{meas}^{ref}(\mathbf{q}) = \text{sinc}\left(\frac{q_x w_x}{2\pi}\right) \text{sinc}\left(\frac{q_y w_y}{2\pi}\right). \tag{B.7}$$

As discussed in the Chapter 2 main text, the size of the illumination window of the reference pixels can be adjusted to compensate for the transmission penalty introduced by the metasurface in the plasmonic devices. In particular, for the metasurface of device A, the $\mathbf{q}$ = 0 transmission for p-polarized light is approximately 45%. At the same time, s-polarized light is almost completed blocked, so that the overall transmission penalty under natural unpolarized illumination is 45%/2 = 23%. Therefore, in order to equate the $\mathbf{q}$ = 0 signals of the reference and plasmonic pixels so that $\Delta T_{tot}(\mathbf{q} = 0) = 0$, we set w_x = 0.23w and w_y = w, where w = 23.8 μm is the size of the plasmonic devices described in the main text. This arrangement has the additional advantages of increased spatial resolution in the x direction, improved noise cancellation upon signal subtraction, and enhanced bandwidth of $T_{meas}^{ref}(\mathbf{q})$ along the q_x direction where filtering takes place. The simulation results plotted in Figure 2.6 were computed with this approach.

B.4 MATLAB code for OTF

For simplicity, here I attach the MATLAB code to calculate $T_+(\mathbf{q})$, and the rest of the incoherent imaging simulation is straightforward.

```
%This is a code to calculate the OTF based on given response maps.
%The code only considers the slits to be on the RIGHT side

%Input: 1. 'farfield' matrix - should be odd number matrix size
    % or 2. 'data' - should be in the polar coordinates (rho*phi)
    % or 3. 'tk_plus'
    % notice input can be either E-field or intensity

    %Notice: the lens imaging system considered here is a single-lens system
    %all dimension units: microns
    % ! if pixel_effect==1, it's the checkerboard arrangement not the
    %compact stacking arrangement


%% input arguments:

polar = 0;          %whether the input data is in the polar coordinates
                %remember that if chosen "polar", "input_type = 2"
input_type = 1;     %1-(simulations)tk_plus and tk_minus data is complex numbers;
                %2-(measurements)input data is intensity.

%choices: 'plasmonic1', 'regular2', 'plasmonic2', 'regular2'
%       and 'nonzero1', 'nonzero2'
%       'nonzero' means the detection peak is not at 0 degree.
%       E.g. dev30, two peaks at +-4 degrees
pixel_type = 'plasmonic1';
                %1 means 1 detectors each pixel
                %2 means 2 detectors each pixel, vertically assembled
                %*effective only when 'pixel_effect'=1
approx = 1;         %if to use the approximation of I_meas

pixel_effect = 0;   %if to include the pixel-size blurring effect
pixel_w = 25.122;   %pixel size
reduction = 4.4;    %the reduction factor of the regular pixel size in x-direction
                %it only applies to 'regular2' pixel type
slits_w = 3.162;    %the space taken by the slits
d = (pixel_w-slits_w/2)/2;
f_num = 3.8;        %f-number of the lens system at image
f_num_ent = 5.5;    %f-number of the object
                %the FOV will be calculated based on f_num
```

```
alpha = 0;        %phase slope of t(k) plus in um. Should be positive.
mag = 1;          %magnification(in this code. Can also be implemented
              % in the incohere_light.m code
interp_num = 2;     %degree of refinery of tk with linear interp method
data_size = 801;    %original farfield plot size from FDTD
              %this can also be used to scale the "interpolation"
              %size
Delta = 0;        %determined transverse coherent length before the lens
              %if incident light is completely incoherent, Delta=0.
              %if Delta to be calculated from FOV, set 0.
peak = 0.38;
lambda = 1.55;

symmetry = 2;      %spin symmetry of the OTF.
rotationN = 1;     %rotate number (0-360/symmetry degree for angle phi)


%% preparing k-space matrices
k0 = 2*pi/lambda;
data_radius = (data_size-1)/2;
theta = atand(1/2/f_num);
if ~Delta
   FOV = 1/f_num_ent;
   Delta = 0.432*lambda/FOV;
end
data_size_1 = (data_size-1)*interp_num+1;
data_radius_1 = interp_num*data_radius;

%intensity distributions contain twice the maximum spatical frequency as
%the original complex-number objects.

     qmax = 2*sind(theta)*k0;
     radius_crop_1 = ceil(qmax/k0*data_radius_1);

size_crop_1 = radius_crop_1*2+1;       %interpolated size of qx qy
tk_size_1 = data_size_1+2*radius_crop_1;

kx = linspace(-1,1,data_size)*k0;      %k-space of initial 'farfield'
ky = kx;
[KX,KY] = meshgrid(kx,ky);
K = sqrt(KX.^2+KY.^2);
kx_1 = linspace(-1,1,data_size_1)*k0;   %k-space for interpolated 'farfield'
ky_1 = kx_1;
[KX_1,KY_1] = meshgrid(kx_1,ky_1);
kx_ext_1 = linspace(-qmax-k0,qmax+k0,tk_size_1);
   %k-space for extended tk_1
ky_ext_1 = kx_ext_1;
```

```
[KX_ext_1,KY_ext_1] = meshgrid(kx_ext_1,ky_ext_1);

qx = linspace(-qmax,qmax,size_crop_1);
qy = qx;
[QX,QY] = meshgrid(qx,qy);
Q = sqrt(QX.^2+QY.^2);


%% mapping 'data' to Cartesian coordinates 'farfield'
if polar==1
  %input: 'data', rho_size*phi_size
  [rho_size,phi_size] = size(data);
  %data = data(:,1:(phi_size-1));
  rho = linspace(0,90,91);
  rho_size = 91;
    %extend rho to integer number of 90 degrees
  phi = linspace(0,360,phi_size);
  [RHO,PHI] = meshgrid(rho,phi);
  data(rho_size,phi_size+1) = 0;  %extend 'data' with 0s
  data = data(:,1:phi_size);
  KX_data = sind(RHO).*cosd(PHI)*k0;
  KY_data = sind(RHO).*sind(PHI)*k0;
  farfield = griddata(KX_data,KY_data,data',KX,KY);
  farfield(isnan(farfield))=0;
  farfield = farfield/max(max(farfield))*peak;
  clear KX_data KY_data;
else
  switch input_type
    case 1
      tk_plus = tk_plus/max(max(tk_plus))*sqrt(peak);
    case 2
      farfield = farfield/max(max(farfield))*peak;
  end
end


%% calculations of tk into T
%expand matrix size to allow convolution:
a = radius_crop_1+1;
b = radius_crop_1+data_size_1;
%------------------------------------------------------------------------
if strcmp(pixel_type,'plasmonic1')||strcmp(pixel_type,'plasmonic2')
  tk_plus1 = zeros(tk_size_1,tk_size_1);
  %application of input_types:
  if input_type==2
    %divided by 2 because at small angle, 2 peaks overlap
    %in this part of the code we consider the contribution of one peak
```

```
    %this only applies to 0-degree devices!!!
    tk_plus1(a:b,a:b) = interp2(KX,KY,sqrt(farfield/2),KX_1,KY_1);
    tk_plus1(a:b,a:b) = tk_plus1(a:b,a:b).*exp(-1i.*alpha.*KX_1);
    %this is t_plus(k)
  else
    tk_plus1(a:b,a:b) = interp2(KX,KY,tk_plus,KX_1,KY_1);
  end
  % tk_1 = tk_1/max(max(tk_1));
  figure(1);
  subplot(1,2,1)
  imagesc(kx_ext_1,ky_ext_1,abs(tk_plus1));
  xlabel('kx');
  ylabel('ky');
  colormap; axis image;
  subplot(1,2,2)
  imagesc(kx_ext_1,ky_ext_1,angle(tk_plus1));
  xlabel('kx');
  ylabel('ky');
  colormap; axis image;

  T_plus = zeros(size_crop_1,size_crop_1);
  dkx = kx_ext_1(2)-kx_ext_1(1);
  dky = ky_ext_1(2)-ky_ext_1(1);
  tk_plus1(isnan(tk_plus1))=0;

  for i=1:size_crop_1
    for j=1:size_crop_1
      %calculating the OTF, T_plus(q) part:
      Conv = conj(tk_plus1(a:b,a:b)).*tk_plus1(i:(i+data_size_1-1),j:(j+data_size_1-1));
      Expo = exp(-Delta^2*((qx(j)+2*KX_1).^2+(qy(i)+2*KY_1).^2)/4);
      %selectively plot the two terms of OTF to check the correctness
      if (i==radius_crop_1+1)&(j==radius_crop_1+1)
        figure(2)
        s2 = surf(Conv);
        s2.EdgeColor = 'none';
      end
      %OTF:
      T_plus(i,j) = sum(sum(Conv.*Expo*dkx*dky));
    end
  end

  T = T_plus.*exp(1i*QX*d);

end


%-----------------------------------------------------------------------
```

```
if strcmp(pixel_type,'regular1')||strcmp(pixel_type,'regular2')
  T = zeros(size_crop_1,size_crop_1);
  dkx = kx_ext_1(2)-kx_ext_1(1);
  dky = ky_ext_1(2)-ky_ext_1(1);
  tk = zeros(tk_size_1,tk_size_1);
  if input_type==2
    tk(a:b,a:b) = interp2(KX,KY,sqrt(farfield),KX_1,KY_1);
  else
    tk(a:b,a:b) = interp2(KX,KY,farfield,KX_1,KY_1);
  end

  figure(1)
  imagesc(kx_ext_1,ky_ext_1,tk);
  xlabel('kx');
  ylabel('ky');
  colormap; axis image;

  for i=1:size_crop_1
    for j=1:size_crop_1
      Conv = conj(tk(a:b,a:b)).*tk(i:(i+data_size_1-1),j:(j+data_size_1-1));
      Expo = exp(-Delta^2*((qx(j)+2*KX_1).^2+(qy(i)+2*KY_1).^2)/4);
      %selectively plot the two terms of OTF to check the correctness
      if (i==radius_crop_1+1)&(j==radius_crop_1+1)
        figure(2)
        s2 = surf(Conv);
        s2.EdgeColor = 'none';
      end
      %OTF:
      T(i,j) = sum(sum(Conv.*Expo*dkx*dky));
    end
  end

end

%% Plotting:
figure(3);
s6 = surf(QX,QY,abs(T));
s6.EdgeColor = 'none';
xlabel('qx');
ylabel('qy');
set(gca,'fontsize',18)
title('OTF')

temp = f_num*Q./k0;
T_lens = 2/pi*(acos(temp)-temp.*sqrt(1-temp.^2));
T_lens = T_lens.*circ(temp);
clear temp;
```

```
%% Introducing the imaging system and sensor physics:
if approx == 0
   T_optical = T_lens.*T;

   %plotting T_optical:
   figure(4);
   s6 = surf(QX,QY,T_optical);
   s6.EdgeColor = 'none';
   xlabel('qx');
   ylabel('qy');

else

%==================================================================
=====
   %enabling the effects of pixel geometry and lens imaging system
   d = (pixel_w-slits_w/2)/2;
   if pixel_effect == 0
      switch pixel_type
         case 'none'
            T_addon = 1;
         case 'regular1'
            T_addon = sinc(QX*pixel_w/reduction/2/pi).*sinc(QY*pixel_w/2/pi);
         case 'plasmonic1'
            T_addon = sinc(QY*pixel_w/2/pi);
         case 'nonzero1'
            T_addon = sinc(QY*pixel_w/2/pi);
         %need to verify!!!
         case 'regular2'
            T_addon = sinc(pixel_w/reduction*QX/2/pi).*sinc(pixel_w*QY/2/pi);
         case 'plasmonic2'
            T_addon = sinc(pixel_w*QY/2/pi);
         case 'nonzero2'
            T_addon = sinc(pixel_w*QY/2/pi);
      end
   else %pixel_effect == 1
      switch pixel_type
         case 'none'
            T_addon = 1;
         case 'regular1'
            T_addon = (sinc(QX*pixel_w/2/pi)).^2.*(sinc(QY*pixel_w/2/pi)).^2;
         case 'plasmonic1'
            T_addon = sinc(QX*pixel_w/2/pi).*(sinc(QY*pixel_w/2/pi)).^2;
         %need to verify!!!
         case 'regular2'
```

```
            T_addon =
sinc(pixel_w/reduction*QX/2/pi).*sinc(pixel_w*QX/2/pi).*sinc(pixel_w*QY/pi).*sinc(pixel_
w*QY/2/pi);
         case 'plasmonic2'
            T_addon = sinc(pixel_w*QX/2/pi).*sinc(pixel_w*QY/pi).*sinc(pixel_w*QY/2/pi);
      end
   end
   %conclusion:
   T_img = T.*T_addon.*T_lens;

   %----------------------------------------------------------------------
   %plotting T_img:
%    T_img = T_img/max(max(T_img));
   figure(4);
   s6 = surf(QX,QY,abs(T_img));
   s6.EdgeColor = 'none';
   xlabel('qx');
   ylabel('qy');
   title('Cumulative OTF T_img')

   figure(5);
   plot(qx,abs(T_img((size_crop_1+1)/2,:)),'linewidth',1.5);
   xlabel('qx');
   title('line cut');
   grid on;
   set(gca,'fontsize',18);
end

%% plotting
rotate = linspace(0,360/symmetry,rotationN+1);
OTF = zeros(size_crop_1,size_crop_1,rotationN);
for m = 1:rotationN
   figure;
   OTF(:,:,m) = imrotate(abs(T_img),rotate(m), 'crop');
   OTF_m = imagesc(qx,qy,OTF(:,:,m));
   axis image; colormap(jet); colorbar;
   set(gca,'fontsize',18)
   xlabel('qx');
   ylabel('qy');
   mycmap = colormap(gca);
   title(['transfer function for',num2str(rotate(m)),' degree']);

   saveas(OTF_m,['transfer_fuction_',num2str(rotate(m)),'deg.jpg']);
end
```

APPENDIX C

Code for Phase Contrast Imaging Simulations

This code serves as the numerical simulation for phase contrast imaging in Chapter 3 Section 4 under the condition of slits being on the left side of the image sensor. Important inputs includes the response map (2D matrix) in the Cartesian coordinate system, with the two dimensions corresponding to $k_x = [-k_0, k_0]$ and $k_y = [-k_0, k_0]$. The phase object is another critical input, which should be a 2D matrix with each data point being the phase value. To reduce the effect of ripples originating from limited Fourier transforms, an option of padding is added for the object. Due to the discretized nature of numerical simulations, the input object and the projected optical image at the sensor array are sampled with a reasonable rate (4 by 4 points per pixel). This code also implements an analysis to the contrast-to-noise ratio when there is a presence of Gaussian white noise on the sensor output. It calls a function “circ” that generates a circular function.

```
%% This code simulates phase contrast imaging. It determines the cut-off frequency
%% of the simulation by the sampling number of each detector pixel.
%% It assumes the input object is already a magnified version.
%% All matrices are even number size except input farfield_1
%% Assume paraxial approximation: NA = 1/2/f_num.
% method of simulating other orientation directions:
% Rotate the input phase object. Run. Rotate back the results and save.


%% input:
    % 'input_type'=tk: t(k), 'tk_plus', or
    % 'input_type'=T: 'farfield_1'.
    % ! for now this code only works for Cartesian input!
    % ! Please make sure the response map is obtained by LEFT slits!
    % 'object1' is the original phase object, if use manual loading
```

```
autoload = 1;                   % if to automatically load an object matrix
                                % if =0, manually read an image into unit8 format, named
                                % 'object', remember to use rgb2gray.
                                % if =0 and object is double, make sure it's [0,255]
autosave = 0;                   % whether to auto save figures and data
lambda = 1.55;                  % in microns
NA_obj = 0.8;                   % this is the physical limitation of the NA of the system
input_type = 'T';               % type 'tk' or 'T'
d_pixel = 21.796;               % sensor pixel size. Use square pixels.
d_sample = 0.2708;              % original object1 resolution, in microns
N = 888;                        % the number of sample points for the actual object and
image
                                % integer number of object1's size is recommended
sample_n = 4;                   % only used for autoload = 1, for artificial gratings.
padding = 0;                    % to how much extent to expand the canvas of the object
polarized = 1;                  % whether the illumination light is polarized
peak = 0.3018;                  % peak of transmission.
                                % NOTICE: be aware of p/s/non-polarized
alpha = 8;                      % phase slope, set 0 to use simulated t(k)
                                % if slits are on the left, alpha>0.
mag = 40.24;                    % magnification. This is to divide NA_obj
% for response map
% N_rot = 1;                    % rotate number (0-180 degree for angle phi)
SNR = 71.3;                     %pixel SNR at full capacity, in decibles
variation = 0.08;

pp_smooth = 4;                  %smoothing of the pupil function's edge

close all;

%% preparation of coordinate systems
% define Fourier transforms
F = @(x) fftshift(fft2(ifftshift(x)));
Ft = @(x) fftshift(ifft2(ifftshift(x)));

k0 = 2*pi/lambda;

% k-space coordinates for farfield_1:
data_size = length(farfield_1(1,:));
data_radius = (data_size-1)/2;
kx = linspace(-k0,k0,data_size);
ky = kx;
[KX,KY] = meshgrid(kx,ky);

% object padding and sampling:
if autoload
```

```
   object1 = rgb2gray(imread('gratings7.bmp'));
   N_obj = length(object1);
   d_sample = d_pixel*N/mag/N_obj/4;
end
Nsize = N*(1+padding);
Nradius = Nsize/2;
refine = N/length(object1(1,:));   % refinement of sampling rate from the original object
d_obj = d_sample/refine;
L = d_obj*Nsize;                   % object physical size

% k-space coordinates for the 1st lens imaging system:
qmax = pi/d_obj;                   % the cut off frequency of sampling
dq = 2*qmax/Nsize;
qx = linspace(-qmax,qmax-dq,Nsize);
qy = qx;
[QX,QY] = meshgrid(qx,qy);

% sample_n should be an integer.
sample_n = d_pixel/(mag*d_obj)
% if abs(sample_n-round(sample_n))>0.1
%     error('down sampling rate is not an integer.')
% end

% k-space coordinates for the magnified object (2nd lens imaging system):
kmax = qmax/mag;
dk_img = 2*kmax/Nsize;
kx_img = linspace(-kmax,kmax-dk_img,Nsize);
ky_img = kx_img;
[KX_img,KY_img] = meshgrid(kx_img,ky_img);
NA_img = NA_obj/mag
f_num = 1/2/NA_img


%% set input
farfield_1 = farfield_1/max(farfield_1(data_radius+1,:))*peak;
crop = interp2(KX,KY,farfield_1,KX_img,KY_img);

figure(1)
set(gcf,'position',[200 200 1000 500]);
subplot(1,2,1)
imagesc(kx,ky,farfield_1);
colorbar; axis image;
title('response map')
xlabel('kx(1/um)'); ylabel('ky(1/um)');
set(gca,'fontsize',18);
subplot(1,2,2)
imagesc(qx,qy,crop);
```

```
colorbar; axis image;
title('expanded, cropped response map');
xlabel('qx(1/um)'); ylabel('qy(1/um)');
set(gca,'fontsize',18);


% pupil function:
qmax_pupil = NA_obj*k0;
pupil = zeros(Nsize,Nsize);
pupil(QX.^2+QY.^2<(qmax_pupil^2))=1;
pupil = smoothdata(pupil,1,'movmean',pp_smooth);
pupil = smoothdata(pupil,2,'movmean',pp_smooth);


% make crop into tk (that is complex):
% tk is the expanded version to fit (qx,qy) coordinates.
switch input_type
   case 'T'
      tk = sqrt(crop).*exp(1i*alpha*(QX/mag));
   case 'tk'
      tk = crop;
end
if ~polarized
   tk = tk/sqrt(2);
end


% reading input object:
% object0 is the actual nonzero object within the canvas:
if autoload
   object1 = double(object1);
   object0 = imresize(object1, [N N], 'nearest');
   object0(object0<0) = 0;
   object0(object0>255) = 255;
   object0 = smoothdata(object0,2,'Gaussian',30);
   object0 = exp(1i*variation*object0/255); % phase object
else
   object0 = imresize(object1, [N N], 'bilinear');
   object0 = exp(1i*object0);
end

% installing the object0 of N onto the canvas of Nsize:
c = (Nsize-N)/2+1;
d = c+(N-1);
background = mean(angle(object0),'all');
object = ones(Nsize,Nsize)*exp(1i*background);
object(c:d,c:d) = object0;
```

```
% coordinates of the object samples:
x = -L/2:d_obj:(L/2-d_obj);
y = x;
% coordinates of the image:
xvec = -L/2:d_obj:(L/2-d_obj); xvec = xvec*mag;
yvec = xvec;
% coordinates of the down-sampled output:
xx = -L*mag/2:d_pixel:(L*mag/2-d_pixel/2);
yy = -(L*mag/2+d_pixel/2):d_pixel:(L*mag/2-d_pixel/2);

figure(2)
set(gcf,'position',[200 100 1000 500]);
subplot(1,2,1)
imagesc(xvec,yvec,angle(object));
colorbar; axis image;
title('phase of the object')

clear object0 KX_img KY_img KX KY QX QY


%% image phase gradient calculation (in x-direction)
object_2 = zeros(Nsize,Nsize);
object_2(:,2:Nsize) = object(:,1:Nsize-1);
phase_diff = angle(object./object_2);
gradient = phase_diff/(x(2)-x(1));
linecut1 = angle(object(Nradius+1,:));

clear object_2 phase_diff

% calculating phase gradient of the object:
figure(3);
set(gcf,'position',[200 100 800 1000])
subplot(2,1,1);
plot(x,linecut1,'linewidth',2);
grid on;
title('object line cut'); xlabel('space coordinate (um)');
set(gca,'fontsize',18);

% calculating light's local deflection angle of the object:
linecut1_2 = gradient(Nradius+1,:);
linecut2 = asind(linecut1_2/k0);
subplot(2,1,2)
plot(x,linecut2/mag,'linewidth',2);
grid on;
title('object local deflection angle/magnification')
```

```
xlabel('x (um)'); ylabel('deflection (degrees)');
set(gca,'fontsize',18);

clear linecut1_2


%% spatial filtering
F_obj = F(object);
figure(2)
subplot(1,2,2)
imagesc(qx,qy,log(abs(F_obj)));
axis image; colorbar;
title('Fourier component of the object logscale')

CTF = pupil.*tk;

figure;
set(gcf,'position',[200 100 1000 500]);
subplot(1,2,1);
imagesc(qx,qy,abs(tk));
axis image;
xlabel('qx'); ylabel('qy');
set(gca,'fontsize',18);
title(['|t(k)|']);

subplot(1,2,2);
imagesc(qx,qy,abs(CTF));
axis image;
xlabel('qx'); ylabel('qy');
set(gca,'fontsize',18);
title('expanded CTF');

image = abs(Ft(F_obj.*CTF)).^2;


%% pixel sampling and reconstruction
% for debuggin purpose, this section is disabled because no need to look at
% the down-sampled signals.
sample_n = round(sample_n);
switch sample_n
  case 3
    reshaped = image(:,1:3:Nsize-2);
    reconstruct = (reshaped(1:3:(Nsize-2),:)+...
      reshaped(2:3:(Nsize-1),:)+reshaped(3:3:Nsize,:))/3;

  case 4
    reshaped = image(:,1:4:Nsize-3);
```

```
    reconstruct = (reshaped(1:4:(Nsize-3),:)+...
      reshaped(2:4:(Nsize-2),:)+reshaped(3:4:(Nsize-1),:)+reshaped(4:4:Nsize,:))/4;

  case 5
    reshaped = image(:,1:5:Nsize-4);
    reconstruct = (reshaped(1:5:(Nsize-4),:)+...
      reshaped(2:5:(Nsize-3),:)+reshaped(3:5:(Nsize-2),:)+...
      reshaped(4:5:Nsize-1,:)+reshaped(5:5:Nsize,:))/5;
end


%% show calculated response & image
figure(5);
set(gcf,'position',[100,100,1200,600])
subplot(1,2,1);
imagesc(xvec/1000, xvec/1000, image); axis image; colorbar; colormap copper;
xlabel('space coordinate (mm)'); ylabel('space coordinate (mm)');
title(['image intensity']);
set(gca,'fontsize',18);

subplot(1,2,2);
imagesc(xx/1000, yy/1000, reconstruct);axis image; colorbar; colormap copper;
xlabel('space coordinate (mm)'); ylabel('space coordinate (mm)');
title(['measured intensity']);
set(gca,'fontsize',18);

%creating a line cut for the reconstructed image:
figure(6);
set(gcf,'position',[100,100,800,800])
G = Nsize/2+1;
subplot(2,1,1);
linecut3 = image(G,:);
plot(xvec/1000,linecut3,'linewidth',2);
grid on;
title('image line cut'); xlabel('space coordinate (mm)');
set(gca,'fontsize',18);

G = round(length(reconstruct)/2);
subplot(2,1,2);
linecut4 = reconstruct(G,:);
plot(xx/1000,linecut4,'linewidth',2);
grid on;
title('reconstruct line cut'); xlabel('space coordinate (mm)');
set(gca,'fontsize',18);


%% noise analysis with the line cut
```

```
SBR = (max(linecut4)-mean(linecut4))/mean(linecut4);
%NOTICE: this code only works for signal~background with high SNR!!
SNR_real = sqrt(mean(linecut4)/peak)*10^(SNR/20);
SNR_log = 20*log10(SNR_real)

CNR = SBR*SNR_real
CNR_log = 20*log10(CNR)
```

APPENDIX D

Derivation and Code for Quantitative DPC Imaging

In this section we derive Equations (3.3) – (3.5) of the main text, which provide the basis for the quantitative phase reconstruction method reported in Chapter 3. These equations describe the sensor array of Figure 3.6(a), which is partitioned into blocks of four adjacent pixels, each coated with the metasurface of Figure 3.3 oriented along one of four orthogonal directions ($\hat{\mathbf{u}} = \pm\hat{\mathbf{x}}$ or $\pm\hat{\mathbf{y}}$). The phase contrast image recorded by all the pixels with the same metasurface orientation $\hat{\mathbf{u}}$ can be computed as

$$\mathrm{I}_{\hat{\mathbf{u}}}(\mathbf{r}) = \iint \frac{\mathbf{dk}}{(2\pi)^2} \mathrm{T}_{\hat{\mathbf{u}}}(\mathbf{k})\mathrm{E}_{\mathrm{obj}}(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{r}} \iint \frac{\mathbf{dk}'}{(2\pi)^2} \mathrm{T}^*_{\hat{\mathbf{u}}}(\mathbf{k}')\mathrm{E}^*_{\mathrm{obj}}(\mathbf{k}')e^{-i\mathbf{k}'\cdot\mathbf{r}}, \quad \text{(D.1)}$$

where $\mathrm{E}_{\mathrm{obj}}(\mathbf{k})$ is the Fourier transform of the incident optical field at the object plane, and $\mathrm{T}_{\hat{\mathbf{u}}}(\mathbf{k})$ is a transfer function that accounts for the $\mathbf{k}$-dependence of the pixels and the imaging optics. For a pure phase object, the optical field is $\mathrm{E}_{\mathrm{obj}}(\mathbf{r}) = \sqrt{\mathrm{P}}e^{i\varphi(\mathbf{r})}$, where P and $\varphi(\mathbf{r})$ indicate power and phase profile, respectively. In the weak-object approximation, [S3] $\mathrm{E}_{\mathrm{obj}}(\mathbf{r})$ can be linearized with respect to $\varphi(\mathbf{r})$, and its Fourier transform becomes

$$\mathrm{E}_{\mathrm{obj}}(\mathbf{k}) \approx \sqrt{\mathrm{P}}[(2\pi)^2\delta(\mathbf{k}) + i\varphi(\mathbf{k})], \quad \text{(D.2)}$$

where $\delta(\mathbf{k})$ is the delta function and $\varphi(\mathbf{k})$ is the Fourier transform of $\varphi(\mathbf{r})$. By substituting Equation (D.2) into Equation (D.1) and neglecting the term proportional to $\varphi(\mathbf{k})\varphi(\mathbf{k}')$, we find

$$\mathrm{I}_{\hat{\mathbf{u}}}(\mathbf{r}) = \mathrm{P}\left\{|\mathrm{T}_{\hat{\mathbf{u}}}(0)|^2 + i \iint \frac{\mathbf{dk}}{(2\pi)^2} [\mathrm{T}^*_{\hat{\mathbf{u}}}(0)\mathrm{T}_{\hat{\mathbf{u}}}(\mathbf{k}) - \mathrm{T}_{\hat{\mathbf{u}}}(0)\mathrm{T}^*_{\hat{\mathbf{u}}}(-\mathbf{k})]\varphi(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{r}}\right\}, \quad \text{(D.3)}$$

where we have used $\varphi(\mathbf{k}) = \varphi^*(-\mathbf{k})$ since $\varphi(\mathbf{r})$ is real.

From Equation (3.1) and its related discussion, the transfer function $T_{\hat{\mathbf{u}}}(\mathbf{k})$ for $\hat{\mathbf{u}} = \pm\hat{\mathbf{x}}$ can be expressed as

$$T_{\pm\hat{\mathbf{x}}}(\mathbf{k}) = C_0 e^{\mp i\alpha k_x} \sqrt{\mathcal{R}(\pm k_{x,} k_{y,})}\, t_{lens}(\mathbf{k}), \qquad \text{(D.4)}$$

where $t_{lens}(\mathbf{k})$ is the contribution from the imaging optics in front of the sensor array, $\mathcal{R}(\mathbf{k})$ is the pixel responsivity map for $\hat{\mathbf{u}} = +\hat{\mathbf{x}}$ (see Figure 3), a is the phase slope computed in Appendix A, and C_0 is a normalization factor. Using this expression, Equation (D.3) for $\hat{\mathbf{u}} = \pm\hat{\mathbf{x}}$ becomes

$$I_{\pm\hat{\mathbf{x}}}(\mathbf{r}) = |C_0|^2 \mathcal{R}(0) P \left\{ 1 + \iint \frac{\mathbf{dk}}{(2\pi)^2} H_x(\pm k_{x,} k_{y,}) \varphi(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} \right\}, \qquad \text{(D.5)}$$

with $H_x(\mathbf{k})$ defined as in Equation (3.4), i.e.,

$$H_x(\mathbf{k}) = i t_{lens}(\mathbf{k}) \left\{ \sqrt{\frac{\mathcal{R}(\mathbf{k})}{\mathcal{R}(0)}} - \sqrt{\frac{\mathcal{R}(-\mathbf{k})}{\mathcal{R}(0)}} \right\}. \qquad \text{(D.6)}$$

In deriving Equations (D.5) and (D.6), we have used the fact that $t_{lens}(\mathbf{k})$ is real, even, and unity at $\mathbf{k} = 0$; furthermore, we have neglected the phase term αk_x of Equation (D.4), which is appropriate for objects with sufficiently slowly varying phase profile. Using Equations (D.5) and (D.6) and the symmetry property $H_x(k_{x,} k_{y,}) = -H_x(-k_{x,} k_{y,})$, we can finally compute the normalized image $S_x(\mathbf{r})$ as follows:

$$S_x(\mathbf{r}) \equiv \frac{I_{+\hat{\mathbf{x}}}(\mathbf{r}) - I_{-\hat{\mathbf{x}}}(\mathbf{r})}{I_{+\hat{\mathbf{x}}}(\mathbf{r}) + I_{-\hat{\mathbf{x}}}(\mathbf{r})} = \iint \frac{\mathbf{dk}}{(2\pi)^2} H_x(\mathbf{k}) \varphi(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}}. \qquad \text{(D.7)}$$

Equation (3.3) [i.e., $S_x(\mathbf{k}) = H_x(\mathbf{k})\varphi(\mathbf{k})$] is equivalent to the Fourier transform of this equation. The same derivation can be used to evaluate $H_y(\mathbf{k})$ and $S_y(\mathbf{r})$, leading to the same expressions of Equations (D.6) and (D.7) with the responsivity map $\mathcal{R}(\mathbf{k})$

rotated by 90°.

The object phase distribution $\varphi(\mathbf{r})$ can then be reconstructed from the measured phase contrast images $I_{\pm\hat{x}}(\mathbf{r})$ and $I_{\pm\hat{y}}(\mathbf{r})$ by solving Equation (D.7) and its counterpart with $x \rightarrow y$. To avoid numerical issues associated with noise amplification at the zeros of $H_x(\mathbf{k})$ and $H_y(\mathbf{k})$, instead of direct inversion we use the following regularized least-squares minimization

$$\min \sum_{u=x,y} |S_u(\mathbf{k}) - H_u(\mathbf{k})\varphi(\mathbf{k})|^2 + \alpha_T \, |\varphi(\mathbf{k})|^2, \tag{D.8}$$

where a_T is a regularization parameter. The solution of this minimization problem is known as Tikhonov regularization [S3] and is given as follows:

$$\varphi(\mathbf{r}) = \mathcal{F}^{-1}\left\{\frac{\sum_{u=x,y} H_u^*(\mathbf{k}) S_u(\mathbf{k})}{\sum_{u=x,y} |H_u(\mathbf{k})|^2 + \alpha_T}\right\}, \tag{D.9}$$

which is the same as Equation (3.5) of the main text.

To conduct numerical simulations, I have developed the following MATLAB code for the quantitative DPC imaging with Tikhonov reconstruction method. The input of the code includes the four output images of the sensor array, corresponding to each orientation of the directional detectors. Then, noise is applied, and phase image is reconstructed.

```
% This code takes the phase contrast imaging simulation results of left-
% and right-orientated pixel arrays and reconstruct the phase object.
% The method of this code is differential phase contrast imaging with
% Tikhonov regularization, as described in the notebook and reference
% paper.
% Important: this code only works for sample_n=4.
% Important: the Gaussian white noise applied to the sensor has a uniform
% SNR according to its average background signal, instead of generated
% pixel-by-pixel.
```

```
%% input:
  % input: 'recon_l', 'recon_r', 'recon_t', 'recon_b', phase contrast imaging simulation results;
  %  'response', cropped response with slits on the left, amplitude
  %  'Nsize', 'sample_n'
  %  'd_pixel'
  %  'pupil', lens pupil function.
  %  'qx', 'qy', 'xx', 'yy'

n_alpha = 0.0000005;            %Tikhonov regularization parameter
alpha = 8;                      %phase slope, in microns.
peak = 0.3018;                  %peak transmission of the response map
SNR = 71.3;                     %maximum SNR, in dB
slope = 0.6908;                 %response map line cut slope at normal incidence

F = @(x) fftshift(fft2(ifftshift(x)));
Ft = @(x) fftshift(ifft2(ifftshift(x)));

close all

%% preparation for variables:
Nsize = length(pupil);          %should be even
Nradius = Nsize/2;
Rsize = length(recon_l);        %should be even
DPCsize = Rsize/2;              %should be even
DPCradius = DPCsize/2;          %should be even

xxx = xx(2:2:Rsize);
yyy = yy(2:2:Rsize);

R_plus = response;
R_minus = flip(R_plus,2);
R0 = R_plus(Nradius+1,Nradius+1);


%% calculation:
% generate Gaussian white noise:
SNR_real = sqrt(mean(recon_l,'all')/peak)*10^(SNR/20)
average = mean(recon_l,'all');

noise = normrnd(0,average/SNR_real,Rsize,Rsize);
I_left = recon_l+noise*1;
noise = normrnd(0,average/SNR_real,Rsize,Rsize);
I_right = recon_r+noise*1;
noise = normrnd(0,average/SNR_real,Rsize,Rsize);
I_top = recon_t+noise*1;
noise = normrnd(0,average/SNR_real,Rsize,Rsize);
I_bot = recon_b+noise*1;
```

```
% calculating DPC measurements:
I_DPC_1 = (I_left-I_right)./(I_left+I_right);
I_DPC_2 = (I_top-I_bot)./(I_top+I_bot);
% down-sampling the DPC measurements (assuming smooth variation of object):
I_DPC_1 = I_DPC_1(2:2:Rsize,2:2:Rsize);
I_DPC_2 = I_DPC_2(2:2:Rsize,2:2:Rsize);

% calculate DPC with Tiknonov regularization:
H = 1i*pupil.*(sqrt(R_plus/R0)-sqrt(R_minus/R0));
H_new = H((Nradius+1-DPCsize/2):Nradius+DPCsize/2,(Nradius+1-
DPCsize/2):Nradius+DPCsize/2);

H1 = H_new;
H2 = rot90(H_new,3);            % equals to rotation clockwisely by 90 degrees

phi_Tik =
Ft((conj(H1).*F(I_DPC_1)+conj(H2).*F(I_DPC_2))./(abs(H1).^2+abs(H2).^2+n_alpha));

% calculate its difference from the object:
object_p = angle(object(4:sample_n*2:Nsize-2,3:sample_n*2:Nsize-2));
difference = real(phi_Tik)-object_p;
difference(difference>pi) = difference(difference>pi)-2*pi;
difference(difference<-pi) = difference(difference<-pi)+2*pi;


%% plotting:
figure(1)
imagesc(xxx,yyy,object_p)
phi_Tik = phi_Tik+mean(angle(object),'all')-mean(real(phi_Tik),'all');

figure(2)
set(gcf,'position',[200 200 1000 500])
subplot(1,2,1)
imagesc(xxx,yyy,real(phi_Tik))
title('Tikhonov reconstruction')
axis image; colorbar;
subplot(1,2,2)
imagesc(xxx,yyy,difference)
title('difference from the object')
axis image; colorbar;

figure(3)
% set(gcf,'position',[200 200 1000 500])
% subplot(1,2,1)
% plot(qx_new,imag(H_new(Rradius,:)));
% subplot(1,2,2)
```

```
object_l = angle(object(Nradius+1,:));
plot(xvec,object_l);
hold on
phi_Tik_l = real(phi_Tik(DPCradius,:));
phi_Tik_l = phi_Tik_l-phi_Tik_l(1)+object_l(1);
plot(xxx,phi_Tik_l)
legend('object','reconstructed');

RMS_diff = rms(difference(DPCradius,:))        % line cut
```

APPENDIX E

Failures in Fabrications

E.1 Outcome of Moisture in SiO_2 Layers

Baking >100 degrees is necessary whenever there is a metal deposition step after the photolithography + development process where SiO_2 layer is exposed to water. If moisture is trapped inside SiO_2 layers, it will evaporate in the next heating process (such as PECVD and baking) and let the covering metal layers pop up.

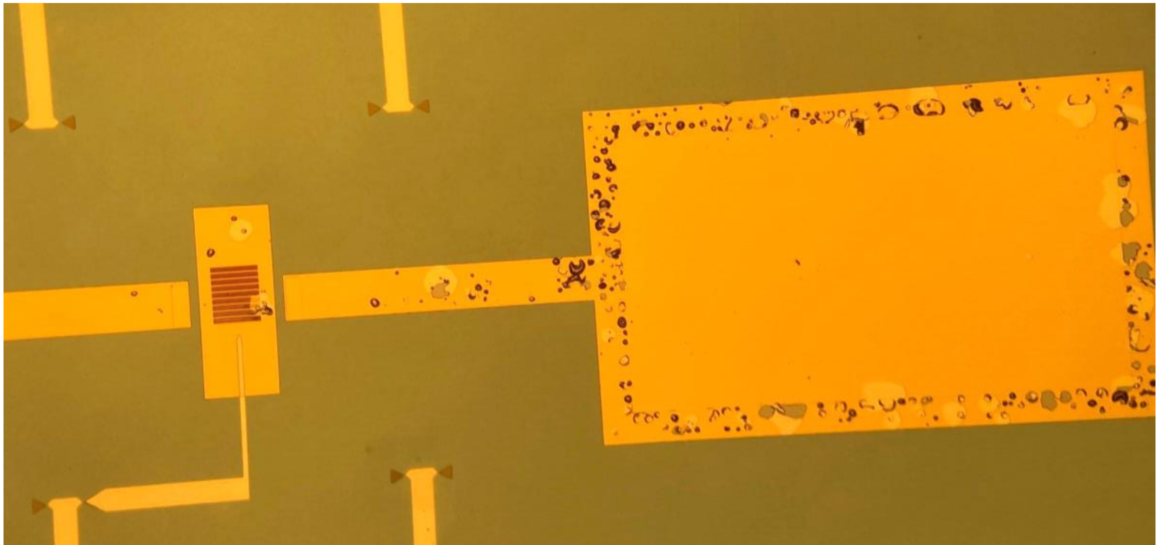

Figure E.1 Photography of a sample with metal films damaged due to inflation of water vapor.

E.2 Attempt of Using HSQ as the Insulation Layer

In the old process, we used HSQ as the insulation layer to prevent short circuit made by the metal contacts and the Ti cover. HSQ forms a conformal layer just by spin coating, and becomes SiO_x insulator compound after exposure to high-energy e-beam. Therefore it has two main advantages: easy to coat, and can be patterned with high precision. This step goes after PECVD to make the spacer layer. The AutoCAD pattern can be found in the same folder.

However, after the whole wafer update, I abandoned this approach based on two reasons. First, when making the second wafer, the top SiO_2 spacer layer as well as the HSQ layer were heavily damaged, completely different from the first trial with a whole wafer, thus proving this method unreliable. Second, HSQ is a very sensitive and expensive chemical (plus additional CNS EBL fees), unlike shared photoresist.

This process is still attractive in some ways so it’s worth reconsidered. An additional advantage compared to PECVD + wet etch process (steps 5 and 6) is obviously the labor and danger: there is no need to use HF and remove the photoresist in the end.

E.3 Attempt of Making the Slits with EBL

Our graduated Ph.D. Leonard Kogos gave up FIB to make the slits and found a new way of making them by double-layer EBL, which is documented in his thesis. However, his difficulty with FIB was due to the incapability of the particular machine, and he acknowledges several advantages of FIB over EBL in his Dissertation.

Double-layer EBL is practiced right after the completion of the bottom SiO_2 layer, creating a continuous Ti + Au film with slits in the device area, followed by a photolithography and wet-etch process to tailor the contacts and center areas. Double-layer EBL requires PMMA as the sacrificial layer and HSQ as the negative e-beam resist. After exposure and development, an RIE process etches down PMMA and forms an under-cut profile, which is very beneficial for metal deposition and lift-off. In addition, the EBL exposure area is quite small, leading to very short exposure and easy lift-off.

However, there are many drawbacks related to this method. Most importantly, the following wet-etch step resulted in huge drop of success rate. Meanwhile, alignment is a problem because there is no high-contrast alignment feature (metal on dielectric) made before this step (metallic alignment marks can be added, but will be an additional complication). Third, the involved RIE process was once very problematic due to the bad directionality of our machine. Lastly, as a factor of inconvenience, the nano-patterns should be already determined before this step, unlike the current version where slits and metasurfaces are made in the very end.

FIB not only bypasses all these obstacles but also offers a straightforward alignment to the metasurface device. But the long milling time and high price of CNS' machine are daunting.

This idea of making slits with EBL is still attractive in some ways but requires an exploration from scratch. There are many new negative e-beam resists on the market such as Allresist AR-N 7520 series that is friendly to lift-off and capable of mix & match processes with photolithography. CNS' EBL system is able to perform local alignment, so a precise alignment of the metasurfaces to the existing slits is also possible.

E.4 Sputtering to Make SiO_2 Layers

In order to make conformal and uniform SiO_2 layers, Angstrom was used to sputter SiO_2 target onto the samples. Sputtering offers very high-quality silica films and the environment is at relatively low temperature. However, Angstrom itself is not very reliable – the vacuum pump often breaks down. Besides, users are asked not to exceed 15% power limit (90W. It used to be 180W limit), further slowing down the already time-consuming deposition. As the target goes through more sputtering, the deposition speed decreases, so calibration runs are required every few months.

BIBLIOGRAPHY

Banerji, S., Meem, M., Majumder, A., Vasquez, F. G., Sensale-Rodriguez, B., & Menon, R. (2019). Imaging with flat optics: metalenses or diffractive lenses? *Optica*, *6*(6), 805.

Bhattarai, K., Silva, S., Song, K., Urbas, A., Lee, S. J., Ku, Z., & Zhou, J. (2017). Metamaterial Perfect Absorber Analyzed by a Meta-cavity Model Consisting of Multilayer Metasurfaces. *Scientific Reports*, *7*(10569).

Bonati, C., Laforest, T., Kunzi, M., & Moser, C. (2020). Phase sensitivity in differential phase contrast microscopy: limits and strategies to improve it. *Optics Express*, *28*(22), 33767–33783.

Bonati, C., Loterie, D., Laforest, T., & Moser, C. (2022). Lock-in incoherent differential phase contrast imaging. *Photonics Research*, *10*(1), 237–247.

Bykov, D. A., Doskolovich, L. L., Bezus, E. A., & Soifer, V. A. (2014). Optical computation of the Laplace operator using phase-shifted Bragg grating. *Optics Express*, *22*(21), 25084.

Chang, J., Sitzmann, V., Dun, X., Heidrich, W., & Wetzstein, G. (2018). Hybrid optical-electronic convolutional neural networks with optimized diffractive optics for image classification. *Scientific Reports*, *8*(1), 12324.

Chen, H., Jayasuriya, S., Yang, J., Stephen, J., Sivaramakrishnan, S., Veeraraghavan, A., & Molnar, A. (2016). ASP Vision: Optically Computing the First Layer of Convolutional Neural Networks Using Angle Sensitive Pixels. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 903.

Chen, W. T., Zhu, A. Y., & Capasso, F. (2020). Flat optics with dispersion-engineered metasurfaces. *Nature Reviews. Materials*, *5*(8), 604–620.

Chen, Y., Zheng, X., Zhang, X., Pan, W., Wang, Z., Li, S., Dong, S., Liu, F., He, Q., Zhou, L., & Sun, S. (2023). Efficient Meta-couplers Squeezing Propagating Light into On-Chip Subwavelength Devices in a Controllable Way. *Nano Letters*, *23*(8), 3326–3333.

Colburn, S., Chu, Y., Shilzerman, E., & Majumdar, A. (2019). Optical frontend for a convolutional neural network. *Applied Optics*, *58*(12), 3179.

Cordaro, A., Kwon, H., Sounas, D., Koenderink, A. F., Alù, A., & Polman, A. (2019). High-Index Dielectric Metasurfaces Performing Mathematical Operations. *Nano Letters*, *19*(12), 8418–8423.

Davis, T. J., Eftekhari, F., Gómez, D. E., & Roberts, A. (2019). Metasurfaces with Asymmetric Optical Transfer Functions for Optical Signal Processing. *Physical Review Letters*, *123*(1), 013901.

Ding, F., Pors, A., & Bozhevolnyi, S. I. (2018). Gradient metasurfaces: a review of fundamentals and applications. *Reports on Progress in Physics*, *81*(2), 026401.

Duparré, J., Dannberg, P., Schreiber, P., Bräuer, A., & Tünnermann, A. (2004). Artificial apposition compound eye fabricated by micro-optics technology. *Applied Optics*, *43*(22), 4303.

Duparré, J., Dannberg, P., Schreiber, P., Bräuer, A., & Tünnermann, A. (2005). Thin compound-eye camera. *Applied Optics*, *44*(15), 2949–2956.

Ebbesen, T. W., Lezec, H. J., Ghaemi, H. F., Thio, T., & Wolff, P. A. (1998). Extraordinary optical transmission through sub-wavelength hole arrays. *Nature*, *391*(6668), 667–669.

Fang, Y., & Ruan, Z. (2018). Optical spatial differentiator for a synthetic three-dimensional optical field. *Optics Letters*, *43*(23), 5893–5896.

Fürhapter, S., Jesacher, A., Bernet, S., & Ritsch-Marte, M. (2005). Spiral phase contrast imaging in microscopy. *Optics Express*, *13*(3), 689.

Gao, L., Qu, Y., Wang, L., & Yu, Z. (2022). Computational spectrometers enabled by nanophotonics and deep learning. *Nanophotonics*, *11*(11), 2507–2529.

Genevet, P., Capasso, F., Aieta, F., Khorasaninejad, M., & Devlin, R. B. (2017). Recent advances in planar optics: from plasmonic to dielectric metasurfaces. *Optica*, *4*(1), 139–152.

Genevet, P., Lin, J., Kats, M. A., & Capasso, F. (2012). Holographic detection of the orbital angular momentum of light with plasmonic photodiodes. *Nature Communications*, *3*(1), 1278.

Gill, P. R., Lee, C., Lee, D.-G., Wang, A., & Molnar, A. (2011). A microscale camera using direct Fourier-domain scene capture. *Optics Letters*, *36*(15), 2949–2951.

Gonzalez, R. C., & Woods, R. E. (2018). *Digital image processing*. Pearson Education Limited.

Goodman, J. W. (2005). *Introduction to Fourier optics*. Roberts and Company Publishers.

Guo, C., Xiao, M., Minkov, M., Shi, Y., & Fan, S. (2018). Photonic crystal slab Laplace operator for image differentiation. *Optica*, *5*(3), 251.

Hamilton, D. K., & Sheppard, C. J. R. (1984). Differential phase contrast in scanning optical microscopy. *Journal of Microscopy*, *133*(1), 27–39.

Hazineh, D. S., Guo, Q., Shi, Z., Huang, Y.-W., Zickler, T., & Capasso, F. (2021). Compact Incoherent Spatial Frequency Filtering Enabled by Metasurface Engineering. *Conference on Lasers and Electro-Optics*. Paper FTu2M.1.

https://doi.org/10.1364/CLEO_QELS.2021.FTu2M.1

Huang, H., Overvig, A., Xu, Y., Malek, S. C., Tsai, C., Alù, A., & Yu, N. (2023). Leaky-wave metasurfaces for integrated photonics. *Nature Nanotechnology*, *18*(6), 580–588.

Huang, L., Chen, X., Bai, B., Tan, Q., Jin, G., Zentgraf, T., & Zhang, S. (2013). Helicity dependent directional surface plasmon polariton excitation using a metasurface with interfacial phase discontinuity. *Light: Science & Applications*, *2*(3), e70.

Huo, P., Zhang, C., Zhu, W., Liu, M., Zhang, S., Zhang, S., Chen, L., Lezec, H. J., Agrawal, A., Lu, Y., & Xu, T. (2020). Photonic Spin-Multiplexing Metasurface for Switchable Spiral Phase Contrast Imaging. *Nano Letters*, *20*(4), 2791–2798.

Iyer, P. P., DeCrescent, R. A., Mohtashami, Y., Lheureux, G., Butakov, N. A., Alhassan, A., Weisbuch, C., Nakamura, S., DenBaars, S. P., & Schuller, J. A. (2020). Unidirectional luminescence from InGaN/GaN quantum-well metasurfaces. *Nature Photonics*, *14*(9), 543–548.

Jayasuriya, S., Sivaramakrishnan, S., Chuang, E., Guruaribam, D., Wang, A., & Molnar, A. (2015). Dual light field and polarization imaging using CMOS diffractive image sensors. *Optics Letters*, *40*(10), 2433–2436.

Ji, A., Song, J.-H., Li, Q., Xu, F., Tsai, C.-T., Tiberio, R. C., Cui, B., Lalanne, P., Kik, P. G., Miller, D. A. B., & Brongersma, M. L. (2022). Quantitative phase contrast imaging with a nonlocal angle-selective metasurface. *Nature Communications*, *13*(1), 7848.

Kim, T., Sridharan, S., & Popescu, G. (2012). Gradient field microscopy of unstained specimens. *Optics Express*, *20*(6), 6737–6737.

Kogos, L. C., Li, Y., Liu, J., Li, Y., Tian, L., & Paiella, R. (2020). Plasmonic ommatidia for lensless compound-eye vision. *Nature Communications*, *11*(1637).

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, *60*(6), 84–90.

Kwon, H., Arbabi, E., Kamali, S. M., Faraji-Dana, M., & Faraon, A. (2020). Single-shot quantitative phase gradient microscopy using a system of multifunctional metasurfaces. *Nature Photonics*, *14*(2), 109–114.

Lam, E. Y. (2015). Computational photography with plenoptic camera and light field capture: tutorial. *Journal of the Optical Society of America A*, *32*(11), 2021.

Li, Z., Kim, M.-H., Wang, C., Han, Z., Shrestha, S., Overvig, A. C., Lu, M., Stein, A., Agarwal, A. M., Lončar, M., & Yu, N. (2017). Controlling propagation and coupling of waveguide modes using phase-gradient metasurfaces. *Nature Nanotechnology*, *12*(7), 675–683.

Ling, T., Boyle, K. C., Goetz, G., Zhou, P., Quan, Y., Alfonso, F. S., Huang, T. W., & Palanker, D. (2018). Full-field interferometric imaging of propagating action potentials. *Light: Science & Applications*, *7*(1), 107.

Liu, J., Wang, H., Kogos, L. C., Li, Y., Li, Y., Tian, L., & Paiella, R. (2022). Optical spatial filtering with plasmonic directional image sensors. *Optics Express*, *30*(16), 29074–29087.

Liu, J., Wang, H., Li, Y., Tian, L., & Paiella, R. (2023). Asymmetric metasurface photodetectors for single-shot quantitative phase imaging. *Nanophotonics*, *12*(17), 3519–3528.

Mehta, S. B., & Sheppard, C. J. R. (2009). Quantitative phase-gradient imaging at high resolution with asymmetric illumination-based differential phase contrast. *Optics Letters*, *34*(13), 1924–1926.

Meng, Y., Chen, Y., Lu, L., Ding, Y., Cusano, A., Fan, J. A., Hu, Q., Wang, K., Xie, Z., Liu, Z., Yang, Y., Liu, Q., Gong, M., Xiao, Q., Sun, S., Zhang, M., Yuan, X., & Ni, X. (2021). Optical meta-waveguides for integrated photonics and beyond. *Light: Science & Applications*, *10*(1), 235.

Mueller, J. P. B., Rubin, N. A., Devlin, R. C., Groever, B., & Capasso, F. (2017). Metasurface Polarization Optics: Independent Phase Control of Arbitrary Orthogonal States of Polarization. *Physical Review Letters*, *118*(11), 113901.

Muminov, B., & Vuong, L. T. (2020). Fourier optical preprocessing in lieu of deep learning. *Optica*, *7*(9), 1079–1088.

Murata, M., Kuroda, R., Fujihara, Y., Otsuka, Y., Shibata, H., Shibaguchi, T., Kamata, Y., Miura, N., Kuriyama, N., & Sugawa, S. (2020). A High Near-Infrared Sensitivity Over 70-dB SNR CMOS Image Sensor With Lateral Overflow Integration Trench Capacitor. *IEEE Transactions on Electron Devices*, *67*(4), 1653–1659.

Paganin, D., & Nugent, K. A. (1998). Noninterferometric Phase Imaging with Partially Coherent Light. *Physical Review Letters*, *80*(12), 2586–2589.

Park, Y., Depeursinge, C., & Popescu, G. (2018). Quantitative phase imaging in biomedicine. *Nature Photonics*, *12*(10), 578–589.

Pors, A., Nielsen, M. G., Bernardin, T., Weeber, J.-C., & Bozhevolnyi, S. I. (2014). Efficient unidirectional polarization-controlled excitation of surface plasmon polaritons. *Light: Science & Applications*, *3*(8), e197.

Pors, A., Nielsen, M. G., & Bozhevolnyi, S. I. (2015). Analog Computing Using Reflective Plasmonic Metasurfaces. *Nano Letters*, *15*(1), 791–797.

Potton, R. J. (2004). Reciprocity in optics. *Reports on Progress in Physics*, *67*(5), 717–754.

Rhodes, W. T. (1980). Incoherent Spatial Filtering. *Optical Engineering*, *19*(3), 323–330.

Saba, A., Tavakol, M. R., Karimi-Khoozani, P., & Khavasi, A. (2018). Two-Dimensional Edge Detection by Guided Mode Resonant Metasurface. *IEEE Photonics Technology Letters*, *30*(9), 853–856.

Saleh, B. E. A., & Teich, M. C. (2007). *Fundamentals of photonics* (2nd ed.). Wiley.

Silva, A., Monticone, F., Castaldi, G., Galdi, V., Alù, A., & Engheta, N. (2014). Performing Mathematical Operations with Metamaterials. *Science*, *343*(6167), 160–163.

Sun, S., He, Q., Xiao, S., Xu, Q., Li, X., & Zhou, L. (2012). Gradient-index meta-surfaces as a bridge linking propagating waves and surface waves. *Nature Materials*, *11*(5), 426–431.

Sun, S., Yang, K.-Y., Wang, C.-M., Juan, T.-K., Chen, W. T., Liao, C. Y., He, Q., Xiao, S., Kung, W.-T., Guo, G.-Y., Zhou, L., & Tsai, D. P. (2012). High-Efficiency Broadband Anomalous Reflection by Gradient Meta-Surfaces. *Nano Letters*, *12*(12), 6223–6229.

Tian, L., Liu, Z., Yeh, L.-H., Chen, M., Zhong, J., & Waller, L. (2015). Computational illumination for high-speed in vitro Fourier ptychographic microscopy. *Optica*, *2*(10), 904–911.

Tian, L., & Waller, L. (2015). Quantitative differential phase contrast imaging in an LED array microscope. *Optics Express*, *23*(9), 11394.

Wang, H., Guo, C., Zhao, Z., & Fan, S. (2020). Compact Incoherent Image Differentiation with Nanophotonic Structures. *ACS Photonics*, *7*(2), 338–343.

Wang, X., Li, Y., Toufanian, R., Kogos, L. C., Dennis, A. M., & Paiella, R. (2020). Geometrically Tunable Beamed Light Emission from a Quantum-Dot Ensemble Near a Gradient Metasurface. *Advanced Optical Materials*, *8*(8), 1901951.

Wang, X., Wang, H., Wang, J., Liu, X., Hao, H., Tan, Y. S., Zhang, Y., Zhang, H., Ding, X., Zhao, W., Wang, Y., Lu, Z., Liu, J., Yang, J. K. W., Tan, J., Li, H., Qiu, C.-W., Hu, G., & Ding, X. (2023). Single-shot isotropic differential interference contrast microscopy. *Nature Communications*, *14*(1).

Wetzstein, G., Ozcan, A., Gigan, S., Fan, S., Englund, D., Soljačić, M., Denz, C., Miller, D. A. B., & Psaltis, D. (2020). Inference in artificial intelligence with deep optics and photonics. *Nature*, *588*(7836), 39–47.

Xie, Y.-Y., Ni, P.-N., Wang, Q.-H., Kan, Q., Briere, G., Chen, P.-P., Zhao, Z.-Z., Delga, A., Ren, H.-R., Chen, H.-D., Xu, C., & Genevet, P. (2020). Metasurface-integrated vertical cavity surface-emitting lasers for programmable directional lasing

emissions. *Nature Nanotechnology*, *15*(2), 125–130.

Yi, S., Xiang, J., Zhou, M., Wu, Z., Yang, L., & Yu, Z. (2021). Angle-based wavefront sensing enabled by the near fields of flat optics. *Nature Communications*, *12*(1), 6002.

Yu, N., & Capasso, F. (2014). Flat optics with designer metasurfaces. *Nature Materials*, *13*(2), 139–150.

Yuan, S., Ma, C., Fetaya, E., Mueller, T., Naveh, D., Zhang, F., & Xia, F. (2023). Geometric deep optical sensing. *Science*, *379*(6637), eade1220.

Zernike, F. (1942). Phase contrast, a new method for the microscopic observation of transparent objects. *Physica*, *9*(7), 686–698.

Zhou, J., Liu, S., Qian, H., Li, Y., Luo, H., Wen, S., Zhou, Z., Guo, G., Shi, B., & Liu, Z. (2020). Metasurface enabled quantum edge detection. *Science Advances*, *6*(51), eabc4385.

Zhou, J., Qian, H., Chen, C.-F., Zhou, J., Li, G., Wu, Q., Luo, H., Wen, S., & Liu, Z. (2019). Optical edge detection based on high-efficiency dielectric metasurface. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(23), 11137–11140.

Zhou, J., Qian, H., Zhou, J., Tang, M., Wu, Q., Lei, M., Luo, H., Wen, S., Chen, S., & Liu, Z. (2020). Two-dimensional optical spatial differentiation and high-contrast imaging. *National Science Review*, *8*(6).

Zhou, R., Edwards, C. J., Arbabi, A., Popescu, G., & Goddard, L. L. (2013). Detecting 20 nm Wide Defects in Large Area Nanopatterns Using Optical Interferometric Microscopy. *Nano Letters*, *13*(8), 3716–3721.

Zhou, Y., Zheng, H., Kravchenko, I. I., & Valentine, J. (2020). Flat optics for image differentiation. *Nature Photonics*, *14*(5), 316–323.

Zhu, T., Zhou, Y., Lou, Y., Ye, H., Qiu, M., Ruan, Z., & Fan, S. (2017). Plasmonic computing of spatial differentiation. *Nature Communications*, *8*(15391).

CURRICULUM VITAE