Computational Interferometric Imaging

Alankar Kotwal

CMU-RI-TR-23-01

December 2022

Robotics Institute
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

Thesis Committee
Ioannis Gkioulekas, chair
Aswin Sankaranarayanan
John Galeotti
Kiriakos Kutulakos (University of Toronto)

This work was supported by NSF Expeditions award 1730147, NSF CAREER award 2047341, the Alfred P. Sloan Research Fellowship, DARPA REVEAL contract HR0011-16-C-0028, ONR DURIP award N00014-16-1-2906, ERC 635537, ISF 1046-14, and the Ollendorff Minerva Center of the Technion.

Abstract

IMAGING systems typically accumulate photons that, as they travel from a light source to a camera, follow multiple different paths and interact with several scene objects. This multi-path accumulation process confounds the information that is available in captured images about the scene, and makes using these images to infer properties of scene objects, such as their shape and material, challenging.

Computational light transport techniques help overcome this multi-path confounding problem, by enabling imaging systems to selectively accumulate only photons that are informative for any given imaging task. Unfortunately, and despite a proliferation of such techniques in the last two decades, they are constrained to operate only under macroscopic settings. This places them out of reach for critical applications requiring microscopic resolutions, such as medical imaging and industrial fabrication.

In this thesis, we change this state of affairs by introducing a new class of techniques that we call computational interferometric imaging: These techniques realize computational light transport capabilities using *optical interferometry*, a technology well-suited for micron-scale applications. We achieve this by either manipulating the properties of illumination used in interferometry setups, or using naturally-available illumination with such properties.

To this end, we develop a theory of interferometry with illumination of general *spatial and temporal coherence*, unifying concepts across incoherent and coherent imaging. We invent and implement a hardware system that realizes such illumination in a programmable and light efficient manner. We then specialize our hardware system to enable two exciting capabilities: (a) fast and robust 3D sensing at micrometer axial and lateral resolutions; and (b) passive, outdoor time-of-flight imaging with sunlight, robust to wind, vibrations and ambient light, factors traditionally considered destructive to interferometry signal. Finally, we provide concrete directions for building faster and more robust extensions of our setups that can be easily deployed in existing interferometry systems in applications ranging from industrial fabrication and inspection, to retinal and cancer imaging.

We hope that this thesis will have a three-fold impact on modern imaging research: (a) providing inspiration for the theoretical investigation of interfer-

ometric and coherent imaging, to demonstrate never-before-seen capabilities; (b) inventing interferometric systems for application areas where interferometry has never been applied; and (c) translating research inspired by this thesis, combined with the development of novel reconstruction techniques, to critical applications in industrial fabrication and biomedical imaging.

Acknowledgments

FIRSTLY, I would like to thank my advisor, Prof. Ioannis Gkioulekas. From high-level feedback to debugging on the optical table, he has always been very resourceful and available. He has actively contributed to my development as a researcher and to me acquiring both the required hardware and software skills to continue forward in my work life. His style of understanding concepts and his skills at presenting them are qualities I wish to emulate.

I'm fortunate to have worked with Prof. Anat Levin, who has been an informal advisor. I greatly benefited from her expertise in optics, computation and optimization. Her critical feedback on theory and experimental results has been invaluable in elevating the quality of the work in this thesis.

I am grateful to my parents, Drs. Shashikant and Manisha Kotwal, and my sister, Ankita Kotwal, for their love and encouragement, and for guiding me onto the path to my PhD. My partner, Aparajita Sahoo, has been with me for all aspects of my PhD life: Her unconditional support throughout my studies has been invaluable in keeping me going, even in moments of doubt.

I'd like to thank my incredible Pittsburgh family: Sudershan Boovaraghavan, Neeha Dev, Prof. Adithya Pediredla, Akhila Kolapalli, Arya Pediredla, Dr. Vishwanath Saragadam, Dr. Mallesh Dasari, Dr. Jayanth Krishna Mogali, Kalyani Tembhe, Samiha Dawalbhakta, Shirsendu Halder, Dr. Ezgi Bakirci, Ashok Kishore, Bhavya Jha, Shubhangi Sathyakumar, Mukesh Thangadurai and Geetha Barani. Over my five years in Pittsburgh, they have made me feel at home, celebrating with my successes and supporting me in times of failure. I'm grateful for the countless laughs, games and dinners I shared with them.

I am thankful to my mentors during my undergraduate studies, Profs. Ajit Rajwade and Suyash Awate, for giving me undergraduate research opportunities. It was through this experience and their philosophies that I understood the value of research and decided to go for a PhD. I would also like to thank Profs. Aswin Sankaranarayanan, Srinivasa Narasimhan and Matt O'Toole for their high-level feedback and guidance over the course of my PhD.

I am also thankful to Brian Hutchison and Jess Butterbaugh for supporting my constant requests: purchasing lab equipment, reimbursing expenses and planning travel. With their expert administrative guidance, my work could proceed smoothly. I would also like to acknowledge the Alfred P. Sloan

Foundation, NSF, DARPA, ONR, and the Ollendorff Minerva Center of the Technion: Without their generous support, this work would not be possible.

Finally, I would like to thank my committee members, Profs. Aswin Sankaranarayanan, Kyros Kutulakos and John Galeotti, for their critical feedback toward making this thesis accurate and complete.

To the prawns with the people sauce

Contents

List of Figures

List of Tables

Chapter 1

Introduction

A camera measures light that leaves an illumination source and travels through a scene of interest. On its way to the camera, light interacts with objects in the scene through mechanisms such as reflection, refraction, and scattering, creating a set of light paths starting at the light source and ending at the camera. A conventional camera integrates contributions from all these light paths indiscriminately. This integration confounds information available in the image about scene properties such as 3D shape and material.

Computational light transport methods aim to mitigate this confounding by selectively measuring the contributions of specific subsets of the paths light travels along in the scene. These subsets are specified in terms of parameters such as path endpoints and length. For example, *transient imaging* techniques measure only those paths with a specific path length, whereas *spatial probing* techniques measure only those paths whose endpoints satisfy specific source-sensor correspondences.

In the last two decades, there has been a proliferation of computational light transport systems based on a large array of technologies, including: projector-camera combinations [35, 40, 50, 50, 52, 53, 70], time-of-flight sensors [24, 27, 55], ultrafast photodiodes [31], streak cameras [69, 71, 72] and single-photon avalanche diodes [17, 54]. These technologies find applications in 3D sensing and material acquisition tasks, and have enabled capabilities previously-thought impossible, such as looking around corners.

Despite this diversity, current computational light transport systems share an important constraint: They are designed for macroscopic imaging, operating at resolutions ranging from a few millimeters to several meters. Resolution here refers not only to the ability of these systems to resolve spatial detail, but more broadly to their ability to

discriminate between light paths. The macroscopic nature of these techniques precludes their application in several critical areas. For example, in industrial fabrication, imaging systems need to assess whether a fabricated part conforms to design specifications at an accuracy of a few micrometers. Likewise, in medical imaging, it is necessary to separate light paths whose lengths differ by no more than a few micrometers, to be able to discern fine structural details below the skin or inside the retina.

A common approach for satisfying the microscopic resolution requirement in the above application domains is to use interferometric imaging techniques. Broadly, these techniques operate by measuring the interference resulting from the superposition of two or more light waves that have traveled along different parts of an optical system. Interference is a consequence of the wave nature of light, and is easily measurable only when working with illumination that is fully- or partially-coherent. As coherence effects are dominant when working at micron scales, interferometric imaging techniques have found widespread use in application areas requiring microscopic imaging resolutions.

1.1 Motivation

Given these advantages of interferometry for micron-scale computational light transport, we are motivated to work towards the following goals and capabilities to unleash its full potential.

Introducing programmable spatial probing to microscopic transient imaging. Traditional implementations of interferometry do very well at discriminating light paths based on micron-scale variations in their length. However, most traditional implementations are very wasteful along the spatial dimension. Interferometer designs used for phase-shifting interferometry and optical coherence tomography are typically designed to physically disallow light scattered from undesired parts of the scene to enter the camera measurement. Such, most commonly *indirect*, light contains valuable insights about the scene: For instance, Liu et al. [41] show that short-range indirect light transport has a lot of information that allows for imaging deep inside human tissue. The contents of the *transmission matrix*, that quantifies full light wave propagation inside the scene, have been shown to be very useful in tasks such as seeing through scattering [4, 30, 44, 57, 58, 62], temporal focusing through scattering media [30, 47, 67] and enhancing energy delivery [66] – however, in measuring pathlength-resolved properties of the scene such as depth or per-pixel transient sequences, most of the transmission matrix is commonly discarded

from the measurements. Thus, there is a lack of coherent interferometric spatio-temporal transmission probing methods akin to the incoherent spatio-temporal transport probing method of O'Toole et al. [51]. The method introduced by Gkioulekas et al. [18], which we build upon, is one step towards coherent interferometric spatio-temporal transmission probing, but is severely limited in the kinds of probing it can perform. In addition, changing probing patterns requires changing the hardware setup.

Using micron-scale probing for inverse scattering. Spatio-temporally probed measurements have been shown to be valuable for *inverse scattering* by Gkioulekas et al. [19]. Inverse scattering is the process of obtaining the pointwise optical scattering properties of an entire 3D volume using intensity measurements taken by illuminating and imaging it at its surface. These measurements are probed versions of the incoherent *light transport matrix* that quantifies the propagation of light intensity in the scene. Probed interferometry can contribute towards microscopic inverse scattering: However, there is a lack of theory connecting probed transmission measurements, that interferometry can provide, to probed light transport.

High lateral resolution with full-field interferometry. Full-field imaging naturally fits within general spatio-probing probing, by not automatically rejecting light from any part of the scene. In addition, performing full-field imaging with a 2D sensor instead of steering optics over the scene decouples the scene size and lateral resolution observed – we can expect to image larger field-of-views with pixel-scale lateral resolutions. We are excited to build and test the effects of this increased lateral resolution, especially in applications such as depth sensing where the benefits of the increased lateral resolution are obvious.

Passive interferometric imaging. Finally, there exist no passive versions of optical interferometry that use naturally-available illumination to image diffuse scenes. The general analysis for probing with illumination we create can be extended to naturally-available illumination, yielding a method for optical interferometry without the need for active illumination. Using the analysis to also manipulate the naturally-available illumination to perform general spatio-temporal probing is an exciting direction for computational interferometric imaging.

1.2 Thesis statement and potential impact

Thesis statement. The goal of this thesis is to take first steps towards implementing the missing interferometric computational light transport capabilities to enhance micron-scale computational imaging. We achieve this objective through four core contributions:

1. *Theoretical foundations.* We establish a theory of interferometric imaging using illumination with given spatio-temporal coherence properties, and show how these properties translate to imaging specific subsets of light paths in a scene, in the context of both coherent light transmission and incoherent light transport (Chapters 2 and 3).
2. *Interferometric computational light transport systems.* We design and implement systems to achieve the required coherence properties. These systems are (a) full-field, allowing for high-resolution imaging and not automatically rejecting useful light, (b) programmable, allowing a user to change the set of selected light paths on the fly, without the need for hardware changes, (c) robust to factors detrimental to interferometry, such as vibrations, noise and ambient light and (d) modular, allowing special cases to be implemented more easily by switching out parts. (Chapter 3).
3. *Task-specific implementations.* We specialize our interferometric systems to perform tasks such as fast, full-field micron-scale 3D shape acquisition (Chapter 4) and outdoor passive micron-scale time-of-flight imaging (Chapter 5). These relatively simple extensions of the above general framework, and the results we obtain with them, demonstrate the power of general spatio-temporal probing, the utility of full-field interferometry setups and test the robustness of our setups in harsh conditions with low SNR.
4. *Establishing concrete future research directions.* Finally, we present, in the future work section, some concrete research directions and interferometer design schematics (one under construction!) that can make computational interferometric imaging faster, easier to implement and closer to being directly applicable to the status quo in interferometry applications.

Potential impact. We envision the potential of the concepts developed in this thesis in the following areas:

1. *Optical biomedical imaging.* Being a non-invasive, contact-less imaging modality,

interferometry in the form of optical coherence tomography is already used for micron-scale soft tissue 3D imaging in ophthalmology, oncology and intravascular monitoring. Our interferometric probing methods are easily incorporated into existing OCT implementations by switching out the light sources – therefore, we envision that the interferometric transport probing capabilities we develop, once deployed into existing optical biomedical imaging setups and combined with inverse scattering, will allow us to acquire higher-quality information about soft tissue structure for diagnosis and continuous monitoring purposes.

2. *Light delivery.* Many applications for light delivery in various fields require knowledge of the transmission matrix of the medium to deliver light through. For example, focusing deep inside living biological tissue using wavefront shaping tailored to the transmission of the tissue allows for, among other capabilities, monitoring and optogenetic modulation of cell activity [61]. The transmission matrix of multimode optical fibers needs to be measured accurately for fast data transmission in communications [60] and multimode fiber-based optical coherence tomography [37]. We envision that fast, closed-loop acquisition of the transmission matrix using methods analogous to the one presented by O'Toole and Kutulakos [49] will allow fast offline and online calibration of these systems.

3. *Surface and sub-surface material inspection.* Another application area that benefits from the contactless nature of interferometry is inspection of critical parts. We will see, in Chapter 4, a prototype system already capable of performing high-resolution micron-scale depth sensing of scenes with a wide variety of material properties. Combined with transient imaging, we expect spatial probing to give us high-resolution transient information for inspection of sub-surface features as well.

4. *Closed-loop precision fabrication.* The addition of the capability to rapidly sense high-resolution depth, as we will show in Chapter 4, gives interferometry the potential to contribute towards closed-loop precision fabrication. We show a result representative of this application in Chapter 4. In addition, there is work on 3D-printing soft biomaterials for tissue grafts [45] to which computational interferometry has significant relevance.

5. *Accessible interferometry.* Finally, we envision that interferometry with passive light sources, of the kind we develop in Chapter 5, combined with some of the acceleration techniques we discuss in the future work, will make interferometry accessible to applications in robotics, manipulation and consumer use.

1.3 Background and related work

We will now establish some mathematical preliminaries that we will use in the rest of this thesis.

Light transport. When light interacts with a scene, it sets off a process of radiative transfer. This process consists of events such as reflection from scene surfaces, refraction through changes in refractive index, and volumetric scattering. This process, assuming the scene consists of media with linear optical properties, is governed by the *radiative transfer equation* (RTE) [6]. The solution to this equation for a given illumination boundary condition characterizes the radiative transfer process pointwise throughout the scene.

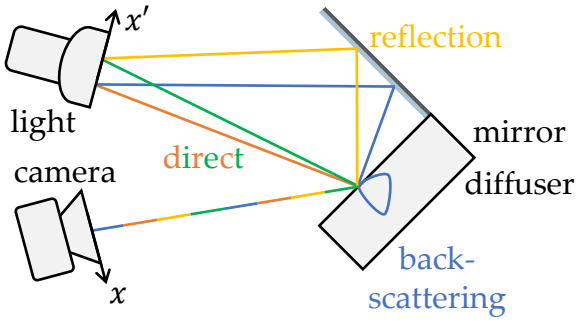


Figure 1.1: **Set of light paths in a scene.** A pixel in a conventional image measures the sum of all lights paths arriving at that pixel. Computational light transport techniques isolate specific subsets of light paths.

The set of all possible solutions to the RTE is characterized by the *Green's function* of the scene $\mathcal{G}(\mathbf{r}, \mathbf{r}')$. The Green's function takes two arguments, both being positions $\mathbf{r}$ and $\mathbf{r}'$ in space. $\mathcal{G}(\mathbf{r}, \mathbf{r}')$ is a solution to the RTE that gives the intensity of light created at $\mathbf{r}$ by a unit light source at $\mathbf{r}'$, upon interacting with the scene. The Green's function can be integrated over all illumination points $\mathbf{r}'$ to get the solution for the given illumination condition, and therefore is the complete description of the spatial interaction of light with the scene.

Consider, in an active imaging scenario, a scene illuminated by a light source and imaged by a camera, as in Figure 1.1. The relevant part of the Green's function for us is then restricted to points x' on the light source, and points x on the camera. This restriction is called the *light transport function* $\mathcal{T}(x, x')$ of the scene, describing, as before, the intensity created at a point x on the camera due to a unit light source at x'. Conventional imaging measures an image, $I(x)$, by integrating light path contributions over all points on the source:

$$I(x) = \int_{x'} \mathcal{T}(x, x') \, \mathrm{d}x'. \tag{1.1}$$

Each $\mathcal{T}(x, x')$ still represents a sum of different paths. As depicted by the blue and orange paths of Figure 1.1, sub-surface scattering and interreflections mean that there

are typically many paths that originate at the same source point and arrive at the same camera pixel, but take different routes through the scene. These paths have different optical lengths τ: Decomposing $\mathcal{T}(x, x')$ along the path length dimension leads to the to the *pathlength-resolved light transport function* $\mathcal{T}(x, x', \tau)$. Then, Equation (1.1) becomes

$$I(x) = \int_{x',\tau} \mathcal{T}(x, x', \tau) \, \mathrm{d}x' \, \mathrm{d}\tau. \tag{1.2}$$

The optical setups we consider are co-located, as shown in Figure 1.2: This means that we use a beamsplitter so that the camera views the scene from the same viewpoint the light source illuminates it. In this co-located setting, light paths contributing to specific subsets of elements of the transport function have names: For instance, the subset $\mathcal{T}(x, x, \tau), \tau \in [0, \infty)$ contains the *diagonal* light paths. The subset of diagonal light paths that bounce only once in the scene are called the *direct-only* light paths. The rest of the diagonal light paths, and light paths contributing to the rest of the transport function, $\mathcal{T}(x, x', \tau), \tau \in [0, \infty), x \neq x'$ are collectively referred to as the *indirect paths*. For an appropriately small ϵ, the subset of indirect light paths contributing to $\mathcal{T}(x, x', \tau), \tau \in [0, \infty), |x - x'| < \epsilon, x \neq x'$ are called *short-range indirect* light paths. These are all names we will come across in the rest of the thesis.

Computational light transport. Computational light transport refers to a set of imaging techniques developed in the last two decades for selectively capturing different components of the light transport in arbitrary scenes. We will restrict ourselves to computational light transport techniques that use, as criteria for filtering light paths, their endpoints and path lengths. We can abstractly describe these techniques as implementing the imaging equation

$$\mathbf{i}(x) = \int_{x',\tau} \mathcal{P}^i(x, x', \tau) \, \mathcal{T}(x, x', \tau) \, \mathrm{d}x' \, \mathrm{d}\tau. \tag{1.3}$$

$\mathcal{P}^i$ is the *probing function*: setting $\mathcal{P}^i(x, x', \tau)$ to zero removes contributions of paths beginning at source point x' and arriving at sensor pixel x after following paths of length τ. In conventional imaging, $\mathcal{P}^i$ is equal to 1 everywhere, and Equation (1.3) reduces to Equation (1.2). Note that discretizing this notation yields the familiar light transport matrix representation as described by O'Toole and Kutulakos [49]. In Chapter 3, we will describe how interferometry with appropriately-designed light sources allows us to perform this probing, and show an implementation of such light sources.

In analogy with the names for the set of light paths, in the co-located case, probing functions that isolate these paths are given names: For instance, a probing function $\mathcal{P}^i(x) = \delta(x - x')$ is called *diagonal probing* and $\mathcal{P}^i(x) = 1 - \delta(x - x')$ is called *indirect-only imaging*. Note that by the definition of probing we have no way to separate the single-bounce, direct light paths from the multi-bounce diagonal light paths: However, in most common diffuse scenes, the single-bounce light path contribution is much higher than multi-bounce light paths, and we often refer to diagonal probing as direct-only imaging.

Optical interferometry. Interferometry refers to techniques that measure the wave-optics phenomenon of interference between one or more light waves [22]. We consider the optical setup shown in Figure 1.2, known as the *full-field Michelson interferometer*. The setup uses a two-dimensional light source and a two-dimensional sensor. This illuminates and images the entire spatial extent of the scene at a time. A beamsplitter divides the light wave emitted by the source into two parts. One part propagates towards the *scene arm* of the interferometer, which contains the scene to be imaged. The other part propagates towards the *reference arm* of the interferometer, which contains some application-specific optical configuration: Most commonly, this is a planar mirror, but we discuss later more general configurations. The reference arm is generally placed on a translation stage that can vary its distance to the beamsplitter to apply a programmable phase shift to the light returning from the reference arm. After reflection, the waves from the two arms recombine at the beamsplitter and reach the camera.

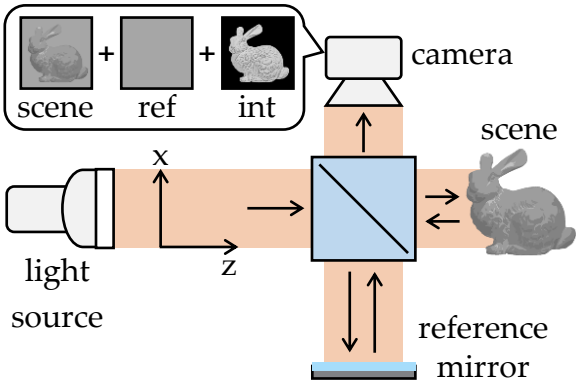


Figure 1.2: **The Michelson interferometer.** A measurement equals the sum of the reference arm image, sample arm image, and an interference component.

We refer to the (x, z) coordinate system in Figure 1.2 for the discussion below. For simplicity, we describe the interferometer in two dimensions, the generalization to three dimensions being straightforward. We use the complex phasor representation for wave fields to write the incoming wave field as $\boldsymbol{u}_i(x, z)$. Upon reflection from the scene and reference, these fields are transformed into $\boldsymbol{u}_s(x, z)$ and $\boldsymbol{u}_r(x, z)$ repectively. The wave

traveling towards the camera is a superposition of the two:

$$\boldsymbol{u}_o\left(x,z\right)=\boldsymbol{u}_s\left(x,z\right)+\boldsymbol{u}_r\left(x,z\right). \tag{1.4}$$

The camera is focused at the $z=0$ plane, and therefore creates a wave field at the sensor plane that is a scaled version of $\boldsymbol{u}_o\left(x,0\right)$. For convenience, we will assume that this scale is 1. Then, the sensor images the intensity $I\left(x\right)$ of this field:

$$\begin{aligned}
I\left(x\right) &= \left|\boldsymbol{u}_o\left(x,0\right)\right|^2 \\
&= \left|\boldsymbol{u}_s\left(x,0\right)+\boldsymbol{u}_r\left(x,0\right)\right|^2 \\
&= \underbrace{\left|\boldsymbol{u}_s\left(x,0\right)\right|^2}_{I_s(x)}+\underbrace{\left|\boldsymbol{u}_r\left(x,0\right)\right|^2}_{I_r(x)}+2\mathrm{Re}\left\{\underbrace{\boldsymbol{u}_s\left(x,0\right)\boldsymbol{u}_r\left(x,0\right)^*}_{\mathcal{C}(x)}\right\} \\
&= \underbrace{I_s\left(x\right)+I_r\left(x\right)}_{I_{DC}(x)}+2\underbrace{\mathrm{Re}\left\{\mathcal{C}\left(x\right)\right\}}_{\mathbf{i}(x)},
\end{aligned} \tag{1.5}$$

where $I_s\left(x\right)$ is the intensity at the camera created by only the scene arm, $I_r\left(x\right)$ is the intensity at the camera created by only the reference arm, and $\mathbf{i}\left(x\right)\equiv\mathrm{Re}\left\{\boldsymbol{u}_s\left(x,0\right)\boldsymbol{u}_r\left(x,0\right)^*\right\}$ is the interference between the two arms. For convenience of calculation, we will represent the interference by the phasor *correlation* of the fields in two arms $\mathcal{C}\left(x\right)\equiv\boldsymbol{u}_s\left(x,0\right)\boldsymbol{u}_r\left(x,0\right)^*$, and use its real part when matching it with physical measurements. In our interference calculations, we will ignore the non-interference term $I_{DC}\left(x\right)\equiv I_s\left(x\right)+I_r\left(x\right)$, and explain how to subtract it from our measurements to isolate the interference in our processing pipelines.

Existing interferometric imaging techniques. A large class of existing high-resolution imaging methods fall within the purview of interferometry [22]. Among the most influential are phase-shifting interferometry [12], optical coherence tomography [18, 25], and synthetic wavelength interferometry [38, 39]. Interferometric methods are typically used for nanometer-scale wavefront sensing and shape acquisition [12], micron-scale subsurface imaging and depth sensing [18, 25, 38, 39], refocusing [10] and non-line-of-sight imaging [73].

Coherence of wave fields. The factor that distinguishes all these variants of interferometry is the properties of the illumination used in the interferometer. These properties are codified in terms of the illumination bandwidth (also referred to as temporal coher-

ence), and second-order time-averaged spatial correlations of the input wave fields (also referred to as mutual intensity, or spatial coherence). When the illumination is monochromatic with the spatial correlations being perfectly 1 in all of space, the illumination is said to be *fully temporally and spatially coherent*, and interferometric imaging with such illumination is called *fully coherent interferometry*. In contrast, when the illumination covers all of the optical spectrum and the correlation is 0 for distinct points, interferometric imaging is said to be *fully incoherent interferometry*

The coherence properties of the illumination are modified to be in between these extreme ends to perform the different kinds of interferometric imaging we mentioned above. As opposed to fully incoherent, we will refer to this form of interferometric imaging as *incoherent interferometry*. In the next chapter, we will derive a theory for incoherent interferometry in the context of the wave-optics analog of light transport, called *light transmission*. In the chapter after that, we will derive the wave optics-based analog of probing with interferometry, and link it to transport probing. We will see how this gives us a method to link all of the different kinds of interferometric imaging, and a way of implementing a general version of these which can emulate their capabilities with no hardware changes. These contributions together form the theoretical basis for this thesis. Specializations of achieved coherence properties to specific situations form the practical contributions and planned future work of this thesis.

Chapter 2

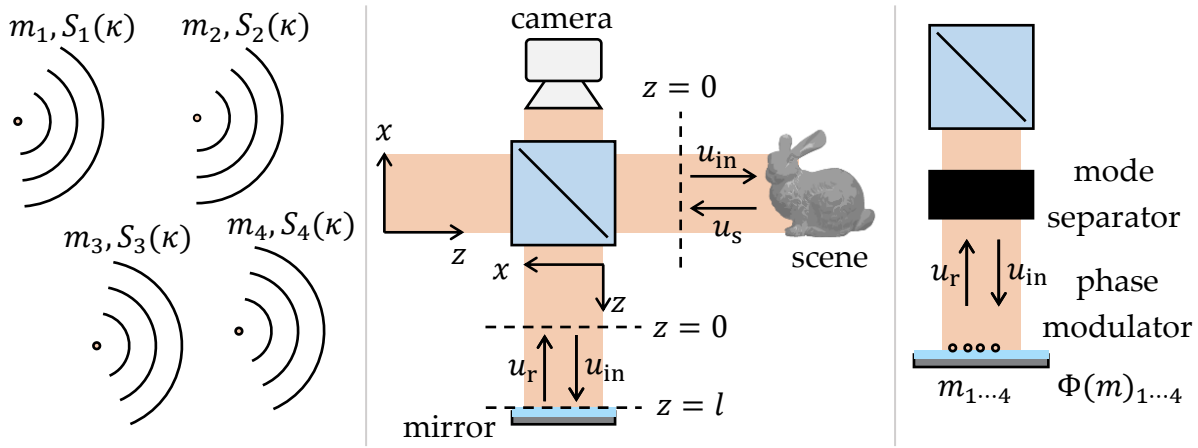
Theory of interferometry with incoherent light

In this chapter, we present a general theory for interferometry using incoming illumination with arbitrary spatio-temporal coherence characteristics. Allowing arbitrary coherence properties significantly complicates the analysis of the interferometric signal: However, we show that with this complexity comes tremendous power and flexibility to measure the optical properties of the sample of interest. In later chapters, we focus on special cases of this general framework, with the coherence properties either programmatically created by us, or given to us by the nature of the incoming illumination.

2.1 The modified Michelson interferometer

In order to perform full-field computational light transport with interferometry, we use an interferometer setup inspired by the Michelson interferometer, but modified to allow for incoherent illumination. In this section, we will discuss the properties of wave fields created in such an interferometer, and then analyze in the next section the interference created at the camera by these wave fields.

Problem setup. To simplify notation, we derive our results in two dimensions, with the extension to the three-dimensional case being straightforward. We use an (x, z) coordinate system, shown in Figure 2.1, where z runs along the optical axis of the illumination lens and the camera. We restrict our derivations to the scalar wave theory, which is sufficient for describing interference effects. We can then describe electromagnetic fields



(a) illumination from transverse modes

(b) full-field Michelson interferometer

(c) generalized reference arm

Figure 2.1: **Michelson interferometer with illumination of general spatio-temporal coherence properties. (a)** Illumination injected into the interferometer consists of a set of transverse modes, each created by a distinct physical light emitter, with their own spectra. **(b)** We use a two-dimensional (x, z) coordinate system for the analysis. **(c)** General reference arm configurations we will consider perform per-transverse-mode phase shifts.

as complex functions that arise as solutions of the Helmholtz equation [20, Section 3.3.1]. For convenience, we set the $z = 0$ at the plane the camera is focused on.

Illumination properties. We assume that the illumination injected into the interferometer is a light wave traveling towards the beamsplitter. Such illumination can be decomposed in two ways. Firstly, we can separate its constituent optical wavelengths λ (wavenumber $\kappa \equiv 2\pi/\lambda$) that correspond to color. Secondly, per optical wavelength, the illumination consists of a collection of transverse modes $m \in \mathcal{M}$, each created by a distinct physical light emitter. The wave field created by the incoming illumination then can be written as

$$\boldsymbol{u}_i(x, z, t) = \int_{\kappa \in [\kappa_{\min}, \kappa_{\max}], m \in \mathcal{M}} A(m)\, S(\kappa; m)\, \boldsymbol{u}_i(x, z, t; \kappa, m)\, \mathrm{d}\kappa\, \mathrm{d}m, \tag{2.1}$$

where $S(\kappa; m)$ is the optical spectrum of the mode m, and $A(m)$ is the amplitude. The utility of this decomposition lies in the *incoherence* of individual components of this decomposition. It can be easily shown that interference between wave fields created by light of different optical wavelengths decreases very rapidly as a function of the integration time of the sensor [5]. Similarly, it can be shown that random phase fluctuations between

distinct physical emitters ensure that the interference between distinct transverse modes decreases very rapidly as a function of the integration time. This can be written in terms of the time-averaged correlations between the terms in Equation (2.1):

$$\left\langle \boldsymbol{u}_i\left(x,z,t;\kappa,m\right),\ \boldsymbol{u}_i\left(x,z,t;\kappa',m'\right)\right\rangle_t = 0 \text{ for all } x,z \text{ if } \kappa \neq \kappa' \text{ and } m \neq m' \tag{2.2}$$

For the purpose of describing interference, a second-order wave effect, we can drop the dependence of the complex functions $\boldsymbol{u}_i\left(x,z,t\right)$ on time to give the complex, time-independent, *phasor* notation $\boldsymbol{u}_i\left(x,z\right)$. In this notation, to derive an expression for the interference between the scene and reference arms caused by this illumination, we calculate it individually for every (κ,m) combination, and then perform an incoherent sum of the interferences over κ and m.

Scene arm. A non-dispersive scene with linear optical properties performs a linear transformation of the input wave field $\boldsymbol{u}_i$. This transformation restricted to a plane with constant z is characterized by the time-resolved *light transmission function* $\mathcal{T}^c\left(x,x',\tau\right)$ of the scene. The light transmission function is analogous to the light transport function $\mathcal{T}\left(x,x',\tau\right)$ we saw in Section 1.3, with $\mathcal{T}^c\left(x,x',\tau\right)e^{-\mathrm{j}c\kappa\tau}$ giving the phasor representation of the wave field at x caused by a unit harmonic wave field of wavenumber κ at x', having been delayed in phase by $\kappa c\tau$. We will evaluate the transmission function at the plane $z=0$ at which the camera is focused. The transformation of $\boldsymbol{u}_i\left(x,0\right)$ to $\boldsymbol{u}_s\left(x,0\right)$ is

$$\boldsymbol{u}_s\left(x,0\right) = \int_{x',\tau} \mathcal{T}^c\left(x,x',\tau\right)\exp\left(-\mathrm{j}c\kappa\tau\right)\ \boldsymbol{u}_i\left(x',0\right)\ \mathrm{d}x'\,\mathrm{d}\tau. \tag{2.3}$$

Reference arm. The reference arm configuration we use is a planar mirror placed on a translation stage that can vary z. We will consider the situation when the reference arm placed at $z=l$. The field reflecting from the reference arm is then a version of $\boldsymbol{u}_i\left(x,0\right)$ propageted by a distance $2l$: $\boldsymbol{u}_r\left(x,0\right)=\boldsymbol{u}_i\left(x,2l\right)$. To account for the effect of this propagation, we decompose the incoming field $\boldsymbol{u}_i\left(x,0\right)$ into its angular spectrum [20, Section 3.10.2]:

$$\boldsymbol{u}_i\left(x,0\right) = \int_{f_x} U\left(f_x;0\right)\exp\left(\mathrm{j}2\pi f_x x\right)\ \mathrm{d}f_x. \tag{2.4}$$

The angular spectrum propagated over a distance $2l$ is:

$$U\left(f_x;2l\right) = U\left(f_x;0\right)\exp\left(-\mathrm{j}2\kappa\beta l\right), \tag{2.5}$$

where $\beta \equiv \sqrt{1-(\lambda f_x)^2}$, yielding the propagated field

$$\boldsymbol{u}_r(x,0) = \boldsymbol{u}_i(x,2l) = \int_{f_x} U(f_x;0)\exp(-\mathrm{j}2\kappa\beta l)\exp(\mathrm{j}2\pi f_x x)\,\mathrm{d}f_x. \tag{2.6}$$

We will make a simplifying paraxial approximation here: $f_x \ll 1/\lambda$, and therefore $\sqrt{1-(\lambda f_x)^2} \approx 1$. We will explain the implications of this assumption at the end of this analysis. This allows us to write

$$\begin{aligned}\boldsymbol{u}_r(x,0) &= \int_{f_x} U(f_x;0)\exp(-\mathrm{j}2\kappa l)\exp(\mathrm{j}2\pi f_x x)\,\mathrm{d}f_x \\ &= \exp(-\mathrm{j}2\kappa l)\int_{f_x} U(f_x;0)\exp(\mathrm{j}2\pi f_x x)\,\mathrm{d}f_x \\ &= \boldsymbol{u}_i(x,0)\exp(-\mathrm{j}2\kappa l).\end{aligned} \tag{2.7}$$

Effectively, the paraxial approximation causes the reference arm to perform an overall phase shift corresponding to propagating a plane wave by a distance $2l$.

2.2 Interferometry with incoherent illumination

We will now derive an expression for the interference component in the framework of the previous section. In the next chapter, we will discuss the ramifications of this derived expression for computational light transport.

We can write the correlation in Equation (1.5) for one κ, m pair in terms of the above fields as follows:

$$\begin{aligned}\mathcal{C}(x) &= \boldsymbol{u}_s(x,0)\,\boldsymbol{u}_r(x,0)^* \\ &= \left[\int_{x',\tau} \mathcal{T}^c(x,x',\tau)\exp(-\mathrm{j}c\kappa\tau)\;\boldsymbol{u}_i(x',0)\;\mathrm{d}x'\,\mathrm{d}\tau\right]\left[(\boldsymbol{u}_i(x,0)\exp(-\mathrm{j}2\kappa l))^*\right] \\ &= \int_{x',\tau} \mathcal{T}^c(x,x',\tau)\exp(-\mathrm{j}\kappa(c\tau-2l))\,\boldsymbol{u}_i(x',0)\,\boldsymbol{u}_i(x,0)^*\;\mathrm{d}x'\,\mathrm{d}\tau.\end{aligned} \tag{2.8}$$

Summed over the optical spectrum and the transverse modes of the illumination, the

interference measures[1]

$$\mathcal{C}(x) = \int_{\kappa,m\in\mathcal{M}} \boldsymbol{u}_s(x,0;\kappa,m)\,\boldsymbol{u}_r(x,0;\kappa,m)^*\,\mathrm{d}\kappa\,\mathrm{d}m \tag{2.9}$$

$$= \int_{x',\tau} \mathcal{T}^c(x,x',\tau) \int_{\kappa,m\in\mathcal{M}} A(m)\,A(m)^*\,\boldsymbol{u}_i(x',0;\kappa,m)\,\boldsymbol{u}_i(x,0;\kappa,m)^* \\ S(\kappa)\exp(-\mathrm{j}\kappa(c\tau-2l))\,\mathrm{d}\kappa\,\mathrm{d}m\,\mathrm{d}x'\,\mathrm{d}\tau \tag{2.10}$$

$$= \int_{x',\tau} \mathcal{T}^c(x,x',\tau) \int_{\kappa,m\in\mathcal{M}} A(m)\,A(m)^*\,\boldsymbol{u}_i(x',0;m)\,\boldsymbol{u}_i(x,0;m)^* \\ S(\kappa)\exp(-\mathrm{j}\kappa(c\tau-2l))\,\mathrm{d}\kappa\,\mathrm{d}m\,\mathrm{d}x'\,\mathrm{d}\tau \tag{2.11}$$

$$= \int_{x',\tau} \mathcal{T}^c(x,x',\tau) \left[\int_\kappa S(\kappa)\exp(-\mathrm{j}\kappa(c\tau-2l))\,\mathrm{d}\kappa\right] \\ \left[\int_{m\in\mathcal{M}} A(m)\,A(m)^*\,\boldsymbol{u}_i(x',0;m)\,\boldsymbol{u}_i(x,0;m)^*\,\mathrm{d}m\right]\mathrm{d}x'\,\mathrm{d}\tau. \tag{2.12}$$

It is also possible to create non-mirror reference arm configurations that add more degrees of freedom to interferometric imaging. One specific generalization we will consider is introducing a mode- and wavelength-dependent phase shifts to the reference field:

$$\boldsymbol{u}_r(x,0;m) = \boldsymbol{u}_i(x,0;m)\exp(-\mathrm{j}2\kappa l)\exp(-\mathrm{j}\Phi(m))\exp(-\mathrm{j}\Psi(\kappa)). \tag{2.13}$$

With this modification, the interference measures

$$\mathcal{C}(x) = \int_{x',\tau} \mathcal{T}^c(x,x',\tau) \left[\int_\kappa S(\kappa)\exp(-\mathrm{j}\kappa(c\tau-2l))\exp(-\mathrm{j}\Psi(\kappa))\,\mathrm{d}\kappa\right] \\ \left[\int_{m\in\mathcal{M}} A(m)\,A(m)^*\,\boldsymbol{u}_i(x',0;m)\,\boldsymbol{u}_i(x,0;m)^*\exp(-\mathrm{j}\Phi(m))\,\mathrm{d}m\right]\mathrm{d}x'\,\mathrm{d}\tau. \tag{2.14}$$

We identify two scene-independent properties here:

$$\mathcal{G}(s) \equiv \int_\kappa S(\kappa)\exp(-\mathrm{j}\kappa s)\exp(-\mathrm{j}\Psi(\kappa))\,\mathrm{d}\kappa \tag{2.15}$$

$$\Gamma(x,x') \equiv \int_{m\in\mathcal{M}} A(m)\,A(m)^*\,\boldsymbol{u}_i(x',0;m)\,\boldsymbol{u}_i(x,0;m)^*\exp(-\mathrm{j}\Phi(m))\,\mathrm{d}m. \tag{2.16}$$

[1]There are simplifying assumptions here: See the paragraph on assumptions below.

When $\Psi(\kappa) = 0$, $\mathcal{G}(s)$ is called the *temporal coherence function* of the illumination, and is the Fourier transform of the optical spectrum. The *mutual intensity function* $\Gamma(x, x')$ specifies the per-wavelength field correlations between points in space, and therfore quantifies the *spatial coherence* of the illumination. With these definitions, we can re-write Equation (2.14):

$$\mathcal{C}(x) = \int_{x',\tau} \mathcal{T}^c(x, x', \tau)\, \mathcal{G}(c\tau - 2l)\, \Gamma(x, x')\, \mathrm{d}x'\, \mathrm{d}\tau. \tag{2.17}$$

In this thesis, we will restrict mutual intensities to depend only on the difference $\epsilon = x - x'$, by considering illumination with space-invariant mutual intensity properties and set $\Psi(\kappa) = 0$. With this in mind, we can re-define the mutual intensity function of points $\Gamma(x, x + \epsilon)$ separated by ϵ as $\Gamma(\epsilon)$:

$$\mathcal{C}(x) = \int_{\epsilon,\tau} \mathcal{T}^c(x, x + \epsilon, \tau)\, \mathcal{G}(c\tau - 2l)\, \Gamma(\epsilon)\, \mathrm{d}x'\, \mathrm{d}\tau. \tag{2.18}$$

Equations (2.15)-(2.18) are the main results of this section. In rest of this thesis, we will analyze these in the context of computational light transport, and derive and implement special cases to perform specific kinds of computational light transport.

Assumptions. We made three major assumptions while performing this derivation. Firstly, we assumed going from Equation (2.6) to Equation (2.7) that $f_x \ll 1/\lambda$, and therefore $\beta = \sqrt{1 - (\lambda f_x)^2} \approx 1$. This amounts to a limitation in the spatial bandwidth of the incoming field from any mode of the illumination: In particular, the field cannot have wavelength-scale variations in the x direction. This requires the incoming modes to be *paraxial*, arising from light emitters close to the optical axis of the interferometer. Specifically, the angular size ϕ subtended by any physical emitter to the optical axis as seen from $z = 0$ must satisfy $\phi \ll \sin^{-1} 1/2\pi \approx 9°$. As we will see in later chapters, the design of our prototypes requires the maximum subtended angle to be as small as possible for optimal signal-to-noise ratios, and we physically limit the maximum subtended angle by closing down the imaging lens aperture on our camera. For the specifications of our imaging system, this limitation does not restrict the kinds of probing we can perform.

In addition, going from Equation (2.10) to Equation (2.11), we dropped the dependence of the fields $\boldsymbol{u}_i(x, 0; \kappa, m)$ on κ. The reasoning behind this is that for a single transverse mode, the entire spectrum arises from the same physical point emitter, which, apart

from propagation phase, creates identical field configurations at $z = 0$ for all κ. We also consider the special case where the spectrum is the same for all modes: $S(\kappa; m) = S(\kappa)$. We leave the general case, *computational interferometry with any light*, where none of these assumptions are fulfilled, to future work. For now, these assumptions still allow for arbitrary spatio-temporal light transport probing with active illumination. The only real case where we need to deal with these is passive interferometry presented in Chapter 5.

2.3 Special cases of incoherent interferometry

We will now illustrate the framework we derived in the previous section by deriving some existing full-field interferometric techniques as special cases of Equation (2.18). We will see the interpretation of these measurements in the context of computational light transport in Section 3.1 of the next chapter.

Phase-shifting interferometry (PSI). Phase-shifting interferometry [12] is version of this general framework performed with fully coherent light, where we set $\Gamma(\cdot) = \mathcal{G}(\cdot) = 1$. With these coherence properties, the interference measurement gives:

$$\mathcal{C}(x) = \int_{\epsilon,\tau} \mathcal{T}^c(x, x+\epsilon, \tau)\, \mathrm{d}\epsilon\, \mathrm{d}\tau. \tag{2.19}$$

Such spatially and temporally coherent illumination is created by manipulating light from a monochromatic, single-mode source such as a laser diode.

Optical coherence tomography (OCT). Optical coherence tomography [25] uses spatially-coherent ($\Gamma(\cdot) = 1$) and temporally-incoherent illumination with a (most commonly Gaussian) spectrum $S(\cdot)$. Using its Fourier transform $\mathcal{G}(\cdot)$ as the temporal coherence function, Equation (2.18) gives:

$$\mathcal{C}(x) = \int_{\epsilon,\tau} \mathcal{T}^c(x, x+\epsilon, \tau)\, \mathcal{G}(c\tau - 2l)\, \mathrm{d}\epsilon\, \mathrm{d}\tau. \tag{2.20}$$

Such illumination is created by manipulating light emitted by a broad-band single-mode emitter such as a superluminescent diode.

Direct-only OCT. In addition to the temporally-incoherent illumination that OCT uses, direct-only OCT [18] uses fully spatially-incoherent illumination with a sharply peaked

mutual intensity: $\Gamma(\epsilon) \approx \delta(\epsilon)$. For such illumination, Equation (2.18) gives:

$$\mathcal{C}(x) = \int_{\epsilon,\tau} \mathcal{T}^c(x, x+\epsilon, \tau)\, \mathcal{G}(c\tau - 2l)\, \delta(\epsilon)\, \mathrm{d}\epsilon\, \mathrm{d}\tau \tag{2.21}$$

$$= \int_{\tau} \mathcal{T}^c(x, x, \tau)\, \mathcal{G}(c\tau - 2l)\, \mathrm{d}\tau. \tag{2.22}$$

Illumination for direct-only OCT is created by manipulating light emitted by a broadband multi-mode emitter such as a light-emitting diode (LED).

Chapter 3

Transmission probing with incoherent interferometry

In this chapter, we discuss the implications of Equation (2.18) for computational light transport. We identify the mutual intensity and temporal coherence functions to be the wave-optics analogs of the spatial and temporal probing patterns used by O'Toole et al. [50, 51] respectively. Whereas O'Toole et al. [50, 51] demonstrate probing the radiometric light transport function of the scene, we theoretically demonstrate that interferometry can probe the coherent light transmission function of the scene with arbitrary, complex probing patterns. Even though, in this thesis, we focus on spatially-invariant mutual intensity leading to spatially-invariant probing patterns, we provide in the future work a direction for generalizing interferometry to probe with spatially-varying patterns. Then, we establish how interferometrically probing the light transmission function also yields a method for probing the light transport function for scenes with highly-randomizing light transmission.

Then, we then present a method for synthesizing illumination with arbitrary spatially-invariant mutual intensity. In this, we focus mostly upon temporally-coherent single-frequency illumination, allowing us to implement *convolutional spatial probing*. We test this method by building a prototype Michelson interferometer modified to generate such illumination. We use this prototype to perform experiments such as visualizing complex fields, capturing direct and global transport components, acquiring light transport, and performing anisotropic descattering. We also combine our technique with full-field optical coherence tomography to perform *spatially-probed transient imaging*.

Potential applications. Measurements of the transmission function (or its discretized version, the transmission matrix) are used in coherent imaging techniques. Exhaustive or partial measurements [58] of the matrix are useful in order to perform tasks such as focusing through scattering [44, 62]. Many techniques performing coherent imaging through scattering media have long prior history in optics and biomedical imaging [4, 29, 57, 67, 68]. Probing with interferometry can be used to acquire these partial measurements through probed patterns, or entire measurements with fast methods such as the one proposed by O'Toole and Kutulakos [49].

3.1 Probing light transmission

In this section, we derive the theoretical framework for interferometric transmission probing that the rest of this thesis is based upon. We will begin with Equation (2.17):

$$\mathcal{C}\left(x\right) = \int_{x',\tau} \mathcal{T}^{c}\left(x, x', \tau\right) \mathcal{G}\left(c\tau - 2l\right) \Gamma\left(x, x'\right) \, \mathrm{d}x' \, \mathrm{d}\tau. \tag{3.1}$$

In analogy with the transport function in Section 1.3, we can interpret the entries of the transmission function $\mathcal{T}^{c}\left(x, x', \tau\right)$ in a ray-optics sense as the coherent sum of contributions of the set of light paths starting at x' and terminating at x, having scattered along paths of length $c\tau$. The effect of the arbitrary coherence properties of the illumination, then, is to create an interference pattern in which contributions of these sets of light paths are weighted by the two illumination-dependent functions. We call this process of weighted summation *probing*. Defining the *probing function* $\mathcal{P}^{c}\left(x, x', s\right) \equiv \mathcal{G}\left(s\right) \Gamma\left(x, x'\right)$, the probed measurement is

$$\mathcal{C}\left(x; \mathcal{P}^{c}\right) = \int_{x',\tau} \mathcal{T}^{c}\left(x, x', \tau\right) \mathcal{P}^{c}\left(x, x', c\tau - 2l\right) \, \mathrm{d}x' \, \mathrm{d}\tau. \tag{3.2}$$

In this interpretation, we can state the first theoretical contribution of this thesis:

Interferometry using illumination with general spatio-temporal coherence properties allows coherent probing of scene light paths grouped by their endpoints and lengths.

Convolutional transmission probing. As we mention before, the mutual intensities we consider in this thesis are spatially-invariant, depending only on the difference

$\epsilon = x - x'$. Our probing functions are then of the form $\mathcal{P}^c(\epsilon, s) \equiv \mathcal{G}(s)\,\Gamma(\epsilon)$. Then, we can write our probed measurements with a change of variables as

$$\mathcal{C}(x; \mathcal{P}^c) = \frac{1}{c} \int_{\epsilon,s} \mathcal{T}^c\left(x, x+\epsilon, {}^{s+2l}/_{c}\right) \mathcal{P}^c(\epsilon, s)\; \mathrm{d}\epsilon\, \mathrm{d}s. \tag{3.3}$$

This amounts to a spatio-temporal convolution of the light transmission function with the probing function. We leave non-convolutional probing with general spatially-varying mutual intensity to future work.

Spatial and temporal probing. The factorization $\mathcal{P}^c(\epsilon, s) = \mathcal{G}(s)\,\Gamma(\epsilon)$ establishes an analogy between the spatial and path length dimensions. This allows us to classify probing into two types: Firstly, we can perform *temporal probing*, where we use a single-mode light source to set $\Gamma(\epsilon) = 1$ and use a non-monochromatic light source with spectrum $S(\kappa)$. This results in a probing function $\mathcal{P}^c(\epsilon, s) = \mathcal{G}(s)$, which is the Fourier transform of the spectrum. Such probing images light paths grouped by only their path lengths, ignoring their endpoints.

A particular case of temporal probing, where we set $\mathcal{G}(s) = \delta(s)$, the Dirac delta function centered at 0, is called *transient imaging*. In this case, the interference with the reference mirror at $z = l$ measures

$$\begin{aligned}\mathcal{C}(x; \mathcal{P}^c) &= \frac{1}{c} \int_{\epsilon,s} \mathcal{T}^c\left(x, x+\epsilon, {}^{s+2l}/_{c}\right) \delta(s)\; \mathrm{d}\epsilon\, \mathrm{d}s \\ &= \frac{1}{c} \int_{\epsilon} \mathcal{T}^c\left(x, x+\epsilon, {}^{2l}/_{c}\right). \end{aligned} \tag{3.4}$$

Transient imaging, therefore, slices the transmission function at ${}^{2l}/_{c}$, allowing us to measure the time-resolved propagation of light in the scene. This measurement is also referred to as the *transient response* of the scene.

Secondly, we can perform *spatial probing*, where we use monochromatic light to set $\mathcal{G}(\kappa) = 1$ and use a multi-mode light source to create a general mutual intensity $\Gamma(\epsilon)$. This results in a probing function $\mathcal{P}^c(\epsilon, s) = \Gamma(\epsilon)$, which images light paths grouped by only their endpoints and ignores their lengths.

Lastly, we can combine the two to perform generalized *spatio-temporal probing* by using a multi-mode non-monochromatic light source with a general mutual intensity $\Gamma(\epsilon)$ and a general temporal coherence function $\mathcal{G}(s)$ to set a general $\mathcal{P}^c(\epsilon, s)$.

Coherence lengths. The strength of convolutional probing achieved by Equation 3.3 is quantified by the widths of the temporal coherence and mutual intensity functions. The width of the mutual intensity function $\Gamma(\epsilon)$ is called the *spatial coherence length* L_Γ of the illumination. If we make the spatial coherence length shorter, the strength of the spatial component of the probing increases, allowing a spatially narrower set of light paths to contribute to the interference. Roughly speaking, the spatial coherence length is inversely proportional to the number of transverse modes in the illumination.

Similarly, the width of the temporal coherence function $\mathcal{G}(s)$ is called the *temporal coherence length* $L_\mathcal{G}$ of the illumination. Making the temporal coherence length shorter increases the strength of the temporal component of the probing increases, narrowing the set of lengths the allowed light paths in the interference travel. Roughly speaking, the temporal coherence length is inversely proportional to the number of wavelengths, or the spectral bandwidth, of the illumination.

Special cases from Section 2.3. We can examine the existing special cases of incoherent interferometry we discussed in Section 2.3 in the context of computational light transport. Firstly, phase-shifting interferometry performs no probing at all, setting $\Gamma(\epsilon) = \mathcal{G}(s) = 1$:

$$\mathcal{C}(x) = \int_{\epsilon,s} \mathcal{T}^c\left(x, x+\epsilon, {}^{s+2l}/_c\right) \, \mathrm{d}\epsilon \, \mathrm{d}s. \tag{3.5}$$

This integral sums up the contributions of paths to sensor point x indiscriminately of their origins and lengths. As we will see in Chapter 4, this indiscriminate summation is often mitigated in its applications by using focusing and steering optics to obtain meaningful information about light paths based on their endpoints and lengths.

Next, optical coherence tomography performs transient imaging by setting a typically sharply peaked Gaussian as its $\mathcal{G}(s)$, and no spatial probing by setting $\Gamma(\epsilon) = 1$. This sums up the contributions of paths to sensor point x indiscriminately of their origins, but gated by their lengths. The strength of the gating is dependent on the temporal coherence length $L_\mathcal{G}$: only light paths of lengths lying in $[2l - {}^{L_\mathcal{G}}/_2, 2l + {}^{L_\mathcal{G}}/_2]$ are allowed into the interference signal. As we will see in Chapter 4, the indiscriminate summation over path origins is often mitigated in its applications by using focusing and steering optics to obtain meaningful information about light paths based on their endpoints.

Last, direct-only OCT performs transient imaging along with spatial probing. As in OCT, only light paths of lengths lying in a window $[2l - {}^{L_\mathcal{G}}/_2, 2l + {}^{L_\mathcal{G}}/_2]$ are allowed into the interference signal, where $L_\mathcal{G}$ is the temporal coherence length. In addition, the only light paths allowed into the interference at sensor pixel x are those that originate in

the window $[x - L_\Gamma/2, x + L_\Gamma/2]$. As we will see further, this property of direct-only OCT makes it very useful for micron-scale depth sensing for sufficiently diffuse scenes.

3.2 Probing light transport

We will now show that, with the framework established in Section 3.1, it is possible to probe the radiometric light transport function for sufficiently diffuse scenes. Whereas probing the transmission function requires the full complex measurement $\mathcal{C}(x)$, probing the transport function requires only its squared magnitude $|\mathcal{C}(x)|^2$.

Relation between transport and transmission. To probe light transport via light transmission, we will first establish an approximate relation between the transmission and transport functions. As in Equation (2.3), the response of a scene with light transmission function $\mathcal{T}^c(x, x', \tau)$ to a wave field $\boldsymbol{u}_i(x, 0)$ is

$$\boldsymbol{u}_s(x, 0) = \int_{x',\tau} \mathcal{T}^c(x, x', \tau) \exp(-\mathrm{j}c\kappa\tau)\, \boldsymbol{u}_i(x', 0)\, \mathrm{d}x'\, \mathrm{d}\tau. \tag{3.6}$$

To examine the relation between the transport and transmission functions, we calculate the output intensity $|\boldsymbol{u}_s(x, 0)|^2$:

$$|\boldsymbol{u}_s(x, 0)|^2 = \int_{x',x'',\tau,\tau'} \mathcal{T}^c(x, x', \tau)\, \mathcal{T}^c(x, x'', \tau')^*\, \boldsymbol{u}_i(x', 0)\, \boldsymbol{u}_i(x'', 0)^* \tag{3.7}$$

$$\exp(-\mathrm{j}c\kappa(\tau - \tau'))\, \mathrm{d}x'\, \mathrm{d}x''\, \mathrm{d}\tau\, \mathrm{d}\tau'. \tag{3.8}$$

The term $\mathcal{T}^c(x, x', \tau)\, \mathcal{T}^c(x, x'', \tau')^*$ in the integral is the correlation between the scattered fields that would be observed at point x when exciting the scene with point sources at point x' and x'', having traveled paths of lengths $c\tau'$ and $c\tau''$ respectively. For scenes with highly-randomizing transmission, this function is approximately band-diagonal, with a very thin non-zero band region [26]:

$$\mathcal{T}^c(x, x', \tau')\, \mathcal{T}^c(x, x'', \tau)^* \approx |\mathcal{T}^c(x, x', \tau)|^2\, \delta(x' - x'')\, \delta(\tau - \tau'). \tag{3.9}$$

The condition of highly-randomizing transmission is typically satisfied by scenes with near-Lambertian surfaces, very rough reflection and transmission, or subsurface multiple

scattering. Then, for such scenes, Equation (3.8) becomes

$$|\boldsymbol{u}_s(x,0)|^2 = \int_{x',\tau} \left|\mathcal{T}^c(x,x',\tau)\right|^2 \left|\boldsymbol{u}_i(x',0)\right|^2 \mathrm{d}x'\,\mathrm{d}\tau. \tag{3.10}$$

The term $|\boldsymbol{u}_i(x',0)|^2$ is the incoming intensity. Therefore, Equation (3.10) describes the transport of intensities, and the coefficient $|\mathcal{T}^c(x,x',\tau)|^2$ is the time-resolved light transport function. Therefore, for scenes with highly-randomizing transmission, the light transport function $\mathcal{T}(x,x',\tau)$ is approximately the squared magnitude of the light transmission function:

$$\mathcal{T}(x,x',\tau) \approx \left|\mathcal{T}^c(x,x',\tau)\right|^2. \tag{3.11}$$

Probing transport via transmission. To probe the transport function with our interferometric imaging setup, we take a look at the squared magnitude of the correlation $\mathcal{C}(x)$ in Equation (2.17).

$$\begin{aligned}|\mathcal{C}(x;\mathcal{P}^c)|^2 = \int_{x',x'',\tau,\tau'} &\mathcal{T}^c(x,x',\tau)\,\mathcal{T}^c(x,x'',\tau')^*\,\Gamma(x,x')\,\Gamma(x,x'')^* \\ &\mathcal{G}(c\tau-2l)\,\mathcal{G}(c\tau'-2l)^*\,\mathrm{d}x'\,\mathrm{d}x''\,\mathrm{d}\tau\,\mathrm{d}\tau'.\end{aligned} \tag{3.12}$$

Again using Equation (3.11), this reduces to

$$|\mathcal{C}(x;\mathcal{P}^c)|^2 = \int_{x',\tau} \left|\mathcal{T}^c(x,x',\tau)\right|^2 \left|\Gamma(x,x')\,\mathcal{G}(c\tau-2l)\right|^2 \mathrm{d}x'\,\mathrm{d}\tau \tag{3.13}$$

$$= \int_{x',\tau} \left|\mathcal{T}^c(x,x',\tau)\right|^2 \left|\mathcal{P}^c(x,x',c\tau-2l)\right|^2 \mathrm{d}x'\,\mathrm{d}\tau \tag{3.14}$$

$$= \int_{x',\tau} \mathcal{T}(x,x',\tau)\;\mathcal{P}^i(x,x',c\tau-2l)\;\mathrm{d}x'\,\mathrm{d}\tau, \tag{3.15}$$

where we define the *transport probing function* $\mathcal{P}^i(x,x',c\tau-2l) \equiv |\mathcal{P}^c(x,x',c\tau-2l)|^2$. Therefore, in scenes with highly-randomizing transmission, the squared magnitude of the correlation $\mathcal{C}(x)$ probes the time-resolved light transport function with the squared magnitude of the probing function $|\mathcal{P}^c(x,x',c\tau-2l)|^2$. Returning to the ray-optics interpretation of the transport function, we can re-write this result as the second theoretical contribution of this thesis:

In scenes with highly-randomizing light transmission, interferometry using illumination with general spatio-temporal coherence properties allows incoherent probing of scene light paths grouped by their endpoints and lengths.

Convolutional transport probing. Analogous to transmission probing, in this thesis, we consider only convolutional transport probing. As in Equation (3.3), we reparameterize $\mathcal{P}^i$ to write our probed measurements as

$$|\mathcal{C}(x;\mathcal{P}^c)|^2 \approx \frac{1}{c}\int_{\epsilon,s} \mathcal{T}\left(x, x+\epsilon, {}^{s+2l}/_{c}\right)\mathcal{P}^i(\epsilon,s)\,\mathrm{d}\epsilon\,\mathrm{d}s. \tag{3.16}$$

As in transmission probing, the factorization $\mathcal{P}^i(\epsilon,s) = |\Gamma(\epsilon)|^2|\mathcal{G}(s)|^2$ allows us to classify transport probing into its spatial and temporal variations.

Relation to transport and probing matrices. Whereas in this thesis we talk of probing *functions*, other probing techniques such as the ones by O'Toole et al. [50, 51] talk of probing *matrices*. We address this by identifying the matrices as discretized versions of the functions:

$$\begin{aligned}\mathbf{T}^t_{nm} &= \mathcal{T}(nr_x, mr_x, tr_t)\\ &\equiv \mathcal{D}_{r_x,r_t}[\mathcal{T}]^t_{nm},\end{aligned}$$

where r_x and r_t are the sampling rates for the spatial and temporal dimensions. A similar discretization can be made for the general non-convolutional probing function[1]:

$$\begin{aligned}\mathbf{P}^t_{nm} &\equiv \mathcal{D}_{r_x,r_t}\left[\mathcal{P}^i\right]^t_{nm}\\ &= \mathcal{P}^i(nr_x, mr_x, tr_t).\end{aligned} \tag{3.17}$$

With these discretizations, we can express the integral in Equation (3.15) as a matrix-vector product summed over the path length dimension:

$$\mathbf{C}^t = \left[\mathbf{T}^t \circ \mathbf{P}^{t'}\right]\mathbf{1} \tag{3.18}$$

$$\mathbf{C} = \sum_t \left[\mathbf{T}^t \circ \mathbf{P}^{t'}\right]\mathbf{1}, \tag{3.19}$$

where $\mathbf{C}^t$ and $\mathbf{C}$ are the time-resolved, and full discretized correlations respectively and t' is shifted to account for the reference mirror translation: $t' = t + \text{integer}\left({}^{2l}/_{r_x}\right)$. A similar expression can be written for probing the transmission matrix $\mathcal{T}^c$ with $\mathcal{P}^c$ by discretizing each of them.

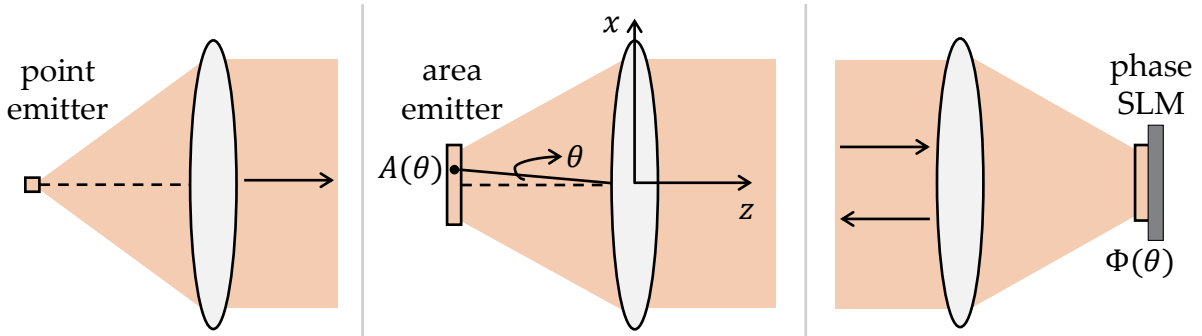
[1]The temporal dimension is treated separately to be consistent with the *matrix* notation of O'Toole et al. [50]. The same equations can be arrived at more compactly by thinking of transport and probing tensors.

Finally, note that when we use a spatially-invariant mutual intensity, and therefore convolutional $\mathcal{P}^c$ or $\mathcal{P}^i$, the entries of the matrices $\mathbf{P}^t_{nm}$ depend only on $m-n$, making them Toeplitz matrices: Therefore, probing using spatially-invariant mutual intensity creates a Toeplitz probing *kernel* that can probe light transport or transmission *functions* with its generated Toeplitz matrix.

3.3 Convolutional spatial probing with coded mutual intensity

We will now present an interferometer design that allows for general convolutional spatial transmission and transport probing. This interferometer design, shown in full in Figure 3.2, is heavily inspired by conventional full-field Michelson interferometry setups.

Illumination design. Full-field interferometry typically uses collimated illumination rising from a point light emitter (such as a laser or superluminescent diode, Figure 3.1(a)) placed in the focal plane of a lens. Since all the light is arises from one point emitter, such illumination is spatially coherent with a constant mutual intensity function $\Gamma(\epsilon) = 1$. We first focus on monochromatic illumination with wavelength λ and wavenumber $\kappa \equiv 2\pi/\lambda$.



(a) conventional illumination

(b) illumination for incoherent interferometry

(c) reference arm phase modulation configuration

Figure 3.1: **Illumination in full-field interferometry. (a)** Conventional Michelson interferometers use a point light source in the focal plane of a collimating lens. **(b)** We replace the point emitter by an area emitter with a controllable amplitude emission profile. **(c)** We change the reference arm configuration in a Michelson interferometer to introduce a varying phase delay across points on the area emitter.

To perform probing, we introduce spatial incoherence into this illumination. To do

so, instead of using a point emitter in the focal plane of the collimating lens, we use an emitter with an extended emission area. Any two physically distinct points on the area emitter are fully incoherent with each other: Therefore, each individual point on the emitter emits one of the transverse modes m in Equation (2.1). In addition, we allow for the mode amplitudes $A(m)$ to be user-programmable. Figure 3.1(b) depicts this modification to the emitter. To allow for the per-mode phase shift in Equation (2.13), we create a custom reference arm configuration shown in Figure 3.1(c).

We will now evaluate Equation (2.16) for this illumination to determine its mutual intensity properties. As shown in Figure 3.1(b), we parameterize the modes by the angular location θ of the light source in the focal plane of the lens. An ideal point emitter located at θ on the focal plane produces a plane wave with amplitude $A(\theta)$ propagating at an angle θ relative to the optical axis. We can express the corresponding field $\boldsymbol{u}_i(x,z)$ as:

$$\boldsymbol{u}_i(x,z;\theta) = A(\theta)\exp\left(-\mathrm{j}\kappa\left(x\sin\theta + z\cos\theta\right)\right), \tag{3.20}$$

We use an area emitter with spatial extent much smaller than the focal length of the collimating lens, causing θ to be small. Therefore, we make the paraxial approximation $\sin\theta = \theta$ and $\cos\theta \approx 1$. We can thus rewrite the plane wave as:

$$\boldsymbol{u}_i(x,z;\theta) = A(\theta)\exp\left(-\mathrm{j}\kappa\left(x\theta + z\right)\right). \tag{3.21}$$

We can then derive the mutual intensity function $\Gamma(x,x')$ of the illumination using Equation (2.16):

$$\begin{aligned}\Gamma(x,x') &= \int_\theta A(\theta)\,A(\theta)^*\left[\exp\left(-\mathrm{j}\kappa\left(x'\theta+z\right)\right)\right]\left[\exp\left(-\mathrm{j}\kappa\left(x\theta+z\right)\right)\right]^*\exp\left(-\mathrm{j}\Phi(\theta)\right)\,\mathrm{d}\theta \\ &= \int_\theta |A(\theta)|^2\exp\left(-\mathrm{j}\kappa\theta\left(x'-x\right)\right)\exp\left(-\mathrm{j}\Phi(\theta)\right)\,\mathrm{d}\theta \end{aligned}\tag{3.22}$$

This mutual intensity funtion is dependent only on the difference $\epsilon \equiv x - x'$, and can be re-written as

$$\Gamma(\epsilon) = \int_\theta |A(\theta)|^2\exp\left(-\mathrm{j}\Phi(\theta)\right)\exp\left(-\mathrm{j}\kappa\epsilon\theta\right)\,\mathrm{d}\theta \tag{3.23}$$

$$= \mathcal{F}\left[|A(\theta)|^2\exp\left(-\mathrm{j}\Phi(\theta)\right)\right](\kappa\epsilon), \tag{3.24}$$

where $\mathcal{F}$ represents the Fourier transform. Thus, the mutual intensity is related, through a Fourier transform, to the spatial intensity distribution of the emitter and the applied per-

mode phase shifts. Then, to create illumination having an arbitrary absolutely-integrable mutual intensity distribution, we use an inverse Fourier transform:

$$H(\theta) = \mathcal{F}^{-1}\left[\Gamma(\epsilon)\right](\theta). \tag{3.25}$$

$H(\theta)$ is typically a complex-valued function. We can then use $A(\theta) = \sqrt{|H(\theta)|}$ and $\Phi(\theta) = \angle H(\theta)$ as the spatial amplitude profile and the phase shifts respectively.

Using monochromatic illumination of the kind depicted in Figure 3.1(a) allows us to perform purely spatial probing:

$$\begin{aligned}\mathcal{C}(x;\mathcal{P}^c) &= \frac{1}{c}\int_{\epsilon,s} \mathcal{T}^c\left(x, x+\epsilon, {}^{s+2l}/_{c}\right)\mathcal{P}^c(\epsilon,s)\,\mathrm{d}\epsilon\,\mathrm{d}s \\ &= \frac{1}{c}\int_{\epsilon,s} \mathcal{T}^c\left(x, x+\epsilon, {}^{s+2l}/_{c}\right)\Gamma(\epsilon)\,\mathrm{d}\epsilon\,\mathrm{d}s\end{aligned} \tag{3.26}$$

where the probing function is $\mathcal{P}^c(\epsilon,s) = \Gamma(\epsilon) = \mathcal{F}\left[|A(\theta)|^2 \exp\left(-\mathrm{j}\Phi(\theta)\right)\right](\kappa\epsilon)$. In our experiments, we will also combine spatial probing with optical coherence tomography using a source with a temporal coherence function given by $\mathcal{G}(s)$, giving the transmission probing function $\mathcal{P}^c(\epsilon,s) = \mathcal{F}\left[|A(\theta)|^2 \exp\left(-\mathrm{j}\Phi(\theta)\right)\right](\kappa\epsilon)\,\mathcal{G}(s)$. The final transmission probing we perform by measuring the complex correlation is of the form

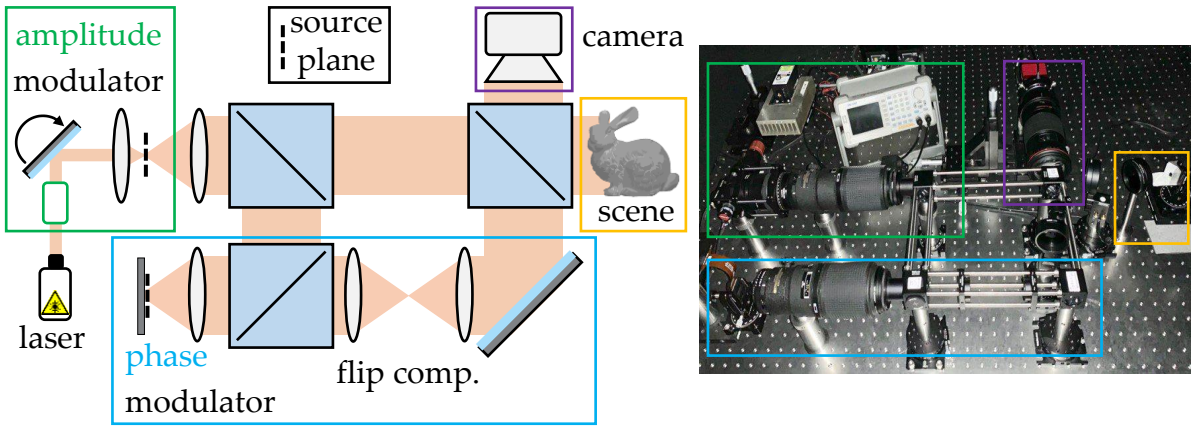
$$\mathcal{C}(x;\mathcal{P}^c) = \frac{1}{c}\int_{\epsilon,s} \mathcal{T}^c\left(x, x+\epsilon, {}^{s+2l}/_{c}\right)\mathcal{F}\left[|A(\theta)|^2 \exp\left(-\mathrm{j}\Phi(\theta)\right)\right](\kappa\epsilon)\,\mathcal{G}(s)\,\mathrm{d}\epsilon\,\mathrm{d}s. \tag{3.27}$$

As in Section 3.2, for sufficiently diffuse scenes, the transport probing we perform by measuring the correlation magnitude is of the form

$$|\mathcal{C}(x;\mathcal{P}^c)|^2 = \frac{1}{c}\int_{\epsilon,s} \mathcal{T}\left(x, x+\epsilon, {}^{s+2l}/_{c}\right)\left|\mathcal{F}\left[|A(\theta)|^2 \exp\left(-\mathrm{j}\Phi(\theta)\right)\right](\kappa\epsilon)\,\mathcal{G}(s)\right|^2 \mathrm{d}\epsilon\,\mathrm{d}s. \tag{3.28}$$

3.4 Implementation

In this section, we discuss how to design and optimize an imaging system for interferometry with coded mutual intensity as in Figure 3.1. Figure 3.2(a) shows a schematic of our design, which we use throughout this section for reference. Figure 3.2(b) shows a photograph of a physical prototype. In the Appendix A, we provide details about constructing, aligning, and operating the setup.



(a) schematic (b) physical prototype

Figure 3.2: **Setup for interferometry with coded mutual intensity. (a)** Schematic of the setup. The setup is a modified Michelson interferometer that includes an illumination component with amplitude modulation (green), and a reference arm with phase modulation (blue). The dotted thick lines indicate the Fourier planes where amplitude and phase modulation take place. **(b)** Photograph of implemented physical prototype. Colors indicate the implementations of the corresponding components in the schematic in (a).

Amplitude modulation. We design an optical configuration for creating a monochromatic area emitter with a specified spatial amplitude profile in the focal plane of the collimating lens. Theoretically, a straightforward way to generate such illumination would be as follows: first, use a uniform monochromatic area emitter whose emission area is larger than the support of $A(\theta)$; second, image this emission area on an amplitude *spatial light modulator* (SLM) applying amplitude modulation $A(\theta)$; finally, collimate the output with a lens. Unfortunately, such an implementation is challenging to realize for two reasons: First, area emitters are typically not sufficiently monochromatic. Second, for sparse amplitude modulation functions $A(\theta)$, this implementation would be very light-inefficient, as most of the source's power would be blocked by the amplitude SLM.

We overcome both challenges by using the setup of Figure 3.2(a): We use a two-dimensional MEMS mirror to steer a collimated coherent laser beam, which is then focused by a scan lens at the focal plane of the main illumination lens. As the direction of the beam incident on the scan lens changes, the focus spot scans the focal plane in a programmable manner, and this scan can take place *within exposure*. Effectively, this scanning scheme corresponds to using time-multiplexing to implement the integration over the incoherent modes θ. To ensure temporal coherence, we use a single-frequency laser, with an estimated temporal coherence length of 20 m. Theoretically, this optical configuration achieves optimal light efficiency, redirecting all of the laser source's power

towards the scene: The steering mirror can be programmed so that the focused point scans only locations θ of the focal plane where $A(\theta)$ is non-zero, and stays at each location for an amount of time proportional to $A(\theta)$. Since this steering over θ causes the incoming plane wave $\boldsymbol{u}_i(x, z; \theta)$ to be *swept* in its angle θ to the optical axis (Equation 3.21), we call this steering mechanism *swept-angle scanning*.

In practice, not all scanning patterns are realizable, both because of acceleration and speed limits imposed by the MEMS mirror, and because the function $A(\theta)$ can be spatially discontinuous, requiring the focused point to instantaneously "jump" from one location θ to another. To address this issue, we place an amplitude *electro-optic modulator* (EOM) between the laser source and the steering mirror, which we synchronize with the mirror: Mirror steering is used to scan only locations θ within the support of $A(\theta)$ (or some superset of this support, as dictated by speed and acceleration limits), and the EOM is used to attenuate the beam at each such location so that the effective overall modulation matches $A(\theta)$. As both the EOM and mirror support MHz operation, this scanning process can take place within exposure. Even though it does not achieve theoretical optimality, the resulting configuration remains significantly more light efficient than the alternative based on an amplitude SLM, and at the same time ensures temporal coherence.

We conclude this discussion with two remarks. First, the configuration we use for amplitude modulation is equivalent to using a laser projector coupled with a scan lens. Unfortunately, the laser diodes in commercial laser projectors have temporal coherence lengths of a few millimeters, making it necessary for us to implement a custom system incorporating a single-longitudinal-mode laser. Second, we can place the above light efficiency considerations within the framework of O'Toole et al. [53]: Using their terminology, when $A(\theta)$ can be realized without an EOM through scanning patterns of the MEMS mirror, the amplitude modulation configuration of Figure 3.2(a) is equivalent to an *impulse projector*. When the EOM is necessary, the configuration is a *redistributive projector*. In both cases, the projector is used to project not directly the probing pattern, but a spatial amplitude distribution derived from its inverse Fourier transform (Equation (3.25)).

Phase modulation. We additionally need to design an optical configuration for implementing the phase modulation Φ. Unlike with amplitude modulation, which can be applied directly on the illumination incident on both arms of the interferometer, the phase modulation needs to be applied only on the reference arm. We achieve this using

the optical configuration shown in Figure 3.2(a): A phase SLM is placed at the focal plane of a lens in the reference arm, and projects a phase modulation pattern equal to Φ. Unlike with amplitude modulation, the use of the phase SLM does not result in light loss, as phase SLMs reflect (most of) the energy incident on them. We note that the combination of the collimating lens with the phase SLM acts as a retroreflector, introducing a reflection of the x coordinate of the beam. We use two additional lenses to cancel this flip ("flip compensation" in Figure 3.2).

Using polychromatic light. We can also use our method with polychromatic illumination, to combine probing with transient imaging. This requires making two modifications to the setup of Figure 3.2: First, we replace the single-frequency laser with a broadband spatially-coherent source. For this, we use a supercontinuum laser, though we could also use a superluminescent diode. Second, we mount the phase SLM and its lens on a translation stage, to be able to scan it at long distances. With these modifications, we capture transient sequences by applying, for each position of the translation stage, the acquisition pipeline described in Section 3.5 for the temporally-coherent case.

3.5 Acquisition pipeline

We now provide details about how to use the optical setup of Section 3.4, in order to capture coherent and incoherent probing measurements.

Implementing a probing pattern. Before capturing measurements with our setup, we need to program it with the desired probing kernel $\mathcal{P}^c$. For this, we compute the inverse Fourier transform of $\mathcal{P}^c$ to obtain the required spatial amplitude profile $A(\theta)$ and phases $\Phi(\theta)$, as in Equation (3.25). These functions are then loaded programmatically on the EOM and MEMS mirror (amplitude) and phase SLM (phase). This calibration process only needs to be performed once per probing measurement.

Capturing probing measurements. Acquiring probing measurements is equivalent to measuring either the complex correlation $\mathcal{C}(x; \mathcal{P}^c)$ (Equation (3.27)) for coherent probing; or its squared amplitude $|\mathcal{C}(x; \mathcal{P}^c)|^2$ (Equation (3.10)) for incoherent probing. From Equation (1.5), we note that an intensity measurement captured at a camera pixel

focused at location x on the plane $z = 0$ equals:

$$I(x) = \underbrace{I_s(x) + I_r(x)}_{\equiv I_{\mathrm{DC}}(x)} + \mathrm{Re}\left\{\mathcal{C}(x; \mathcal{P}^c)\right\}. \tag{3.29}$$

Analogously to continuous-wave time-of-flight cameras [21] and phase-shifting interferometry [11], we estimate the correlation term in a two-step procedure: First, we capture $N \geq 3$ intensity measurements, where for each of them we use the phase SLM to shift the phase modulation by a sub-wavelength amount, $\Phi(\theta) + 2\pi n/N$, $n = 0, \ldots, N$. Second, at each pixel, we fit a sinusoid to these N measurements, to obtain the per-pixel amplitude and phase of the correlation term $\mathcal{C}(x; \mathcal{P}^c)$. For incoherent probing, we directly estimate the squared amplitude $|\mathcal{C}(x; \mathcal{P}^c)|^2$ as:

$$I_{\mathrm{DC}}(x) = \frac{1}{N} \sum_{n=1}^{N} I_n(x), \tag{3.30}$$

$$|\mathcal{C}(x; \mathcal{P}^c)|^2 = \frac{1}{N} \sum_{n=1}^{N} \left(I_n(x) - I_{\mathrm{DC}}(x)\right)^2. \tag{3.31}$$

In practice, we found that we can reliably estimate the correlation term or its squared amplitude using measurements at $N = 10$ sub-wavelength phase shifts. We note that, as phase shifting is performed by the phase SLM and no mechanical parts are involved, these measurements can be captured at a frame rate of 6 Hz. This is limited by the maximum refresh rate of the phase SLM (60 Hz).

Dealing with speckle. In real scenes, the correlation measurements $\mathcal{C}(x; \mathcal{P}^c)$ will contain significant pseudo-random speckle noise. When doing incoherent probing, to eliminate these speckle artifacts, we blur our estimates with a small blur kernel $B(x)$. Therefore, our final estimate for incoherent probing is:

$$|\mathcal{C}(x; \mathcal{P}^c)|^2 * B(x). \tag{3.32}$$

For a detailed analysis of speckle in interferometry, we refer to Gkioulekas et al. [18]. We follow their suggestions on how to set magnification and aperture size to maximize interference contrast. We also use a neutral density filter in the reference arm of our setup (Figure 3.2), to equalize the intensity of the reference and target arms.

3.6 Experiments

We now show results of experiments performed using the prototype described in Section 3.4. Our experiments emphasize probing types that cannot be implemented using previous interferometric techniques [18], or do not have light-efficient implementations using projector-camera techniques [50]. The project website [34] shows additional results, including full video sequences for the experiments combining probing with transient imaging. To facilitate reproducibility, we provide all of our data at the project website.

Coherent probing and mutual intensity visualization. In Figure 3.3, we perform experiments to coherently probe the transmission matrix of a retroreflector scene. A retroreflector has the property that it rotates the incoming field by 180° about its center. Specifically, the transmission function of a retroreflector has the form $\mathcal{T}^c\left(x, x', \tau\right) = \delta\left(x + x'\right)\delta\left(\tau - {}^{z_0}/_c\right)$, where z_0 is the z-coordinate of the retroreflector center and δ is the Dirac delta function. Then, Equation (3.2) yields for a probing function $\mathcal{P}^c\left(\epsilon, s\right) = \Gamma\left(\epsilon\right)$:

$$\begin{aligned}\mathcal{C}\left(x; \mathcal{P}^c\right) &= \frac{1}{c}\int_{\epsilon,s} \mathcal{T}^c\left(x, x+\epsilon, {}^{s+2l}/_c\right)\mathcal{P}^c\left(\epsilon, s\right)\,\mathrm{d}\epsilon\,\mathrm{d}s \\ &= \frac{1}{c}\int_{\epsilon,s} \delta\left(2x+\epsilon\right)\delta\left({}^{s+2l-z_0}/_c\right)\Gamma\left(\epsilon\right)\,\mathrm{d}\epsilon\,\mathrm{d}s \\ &= \frac{\Gamma\left(-2x\right)}{c}\end{aligned} \tag{3.33}$$

Thus, the complex correlation created by probing a retroreflector is exactly the complex mutual intensity function. We image on the camera the real part of the mutual intensity function, which adds a sinusoidal modulation upon the camera image, which appears as fringes. The local amplitude of the fringe pattern gives the mutual intensity function.

In Figure 3.3, we show experiments using a variety of one-dimensional probing functions corresponding to amplitude modulation functions $A\left(\theta\right)$ of different sizes (Figure 3.3(a)), orientations (Figure 3.3(b)), and profiles (Figure 3.3(c)). In the latter case, we implement one-dimensional sources with horizontal emission profiles corresponding to amplitude modulation that is: approximately constant, $A(\theta) \propto 1$; Gaussian, $A(\theta) \propto \exp(-\theta^2/\sigma^2)$ for some standard deviation σ; and the inverse Fourier transform of the Laplacian-of-Gaussian function, $A(\theta) \propto \theta^2\exp(-\theta^2/\sigma^2)$. In all cases, the resulting probing function is constant along the direction that is orthogonal to the orientation of $A\left(\theta\right)$ – for example, when $A\left(\theta\right)$ is horizontal, the probing function is constant along the vertical direction. In Figure 3.3(a), we observe that as we make $A\left(\theta\right)$ narrower, the

support of the probing function becomes wider. In Figure 3.3(b), we observe that as we rotate $A(\theta)$, the probing function rotates as well. Both of these observations correspond to standard properties of the Fourier transform. Finally, in Figure 3.3(c), we observe that the profile of the probing function matches the Fourier transform of $A(\theta)$.

Direct-indirect separation. From the Fourier transform relationship of Equation 3.24, the probing function due to a two-dimensional square amplitude modulation function $A(\theta)$ is a two-dimensional sinc. As the size of $A(\theta)$ increases, the central lobe of the probing function shrinks and the secondary lobes become negligible – that is, the probing function approximates a Dirac delta function. This is equivalent to probing the transmission and transport matrices to set $x = x'$. As explained by O'Toole et al. [50], this is approximately equivalent to measuring only the direct and retroreflecting components of light propagation in a scene. We also subtract the estimated direct component from the scene image to obtain a indirect-enhanced image.

Figure 3.4 shows three example scenes where we measure the direct and indirect components of the incoherent light transport matrix. The first scene uses a small toy cup, resulting in three types of light paths: direct reflections; retroreflective specular, which show up as the strong specularities at the upper and lower parts of the cup; and indirect paths due to diffuse interreflections and sub-surface scattering. The direct-only measurement removes the last two types of paths, making the cup appear metallic. The indirect-enhanced measurement, on the other hand, includes the caustics, interreflections, and translucent effects due to the indirect paths.

The second scene consists of a plastic transparent bead placed between two diffuse walls. The bead faces reflect light and create strong specularities on the walls. Additionally, some of the light from the walls reflects specularly on the bead and towards the camera, resulting in specular reflections on some of the bead faces. All these specular effects are removed from the direct-only measurement, and are prominent in the indirect-enhanced measurement. The indirect-enhanced component additionally shows effects due to the interreflections between the diffuse walls and inside the bead.

The third scene is a gummy bear placed between two diffuse walls. The majority of the light received from the gummy bear is due to subsurface scattering inside it. This light is removed from the direct-only component, resulting in the gummy bear appearing metallic due to surface-only reflections. By contrast, the indirect-enhanced component is dominated by this scattered light.

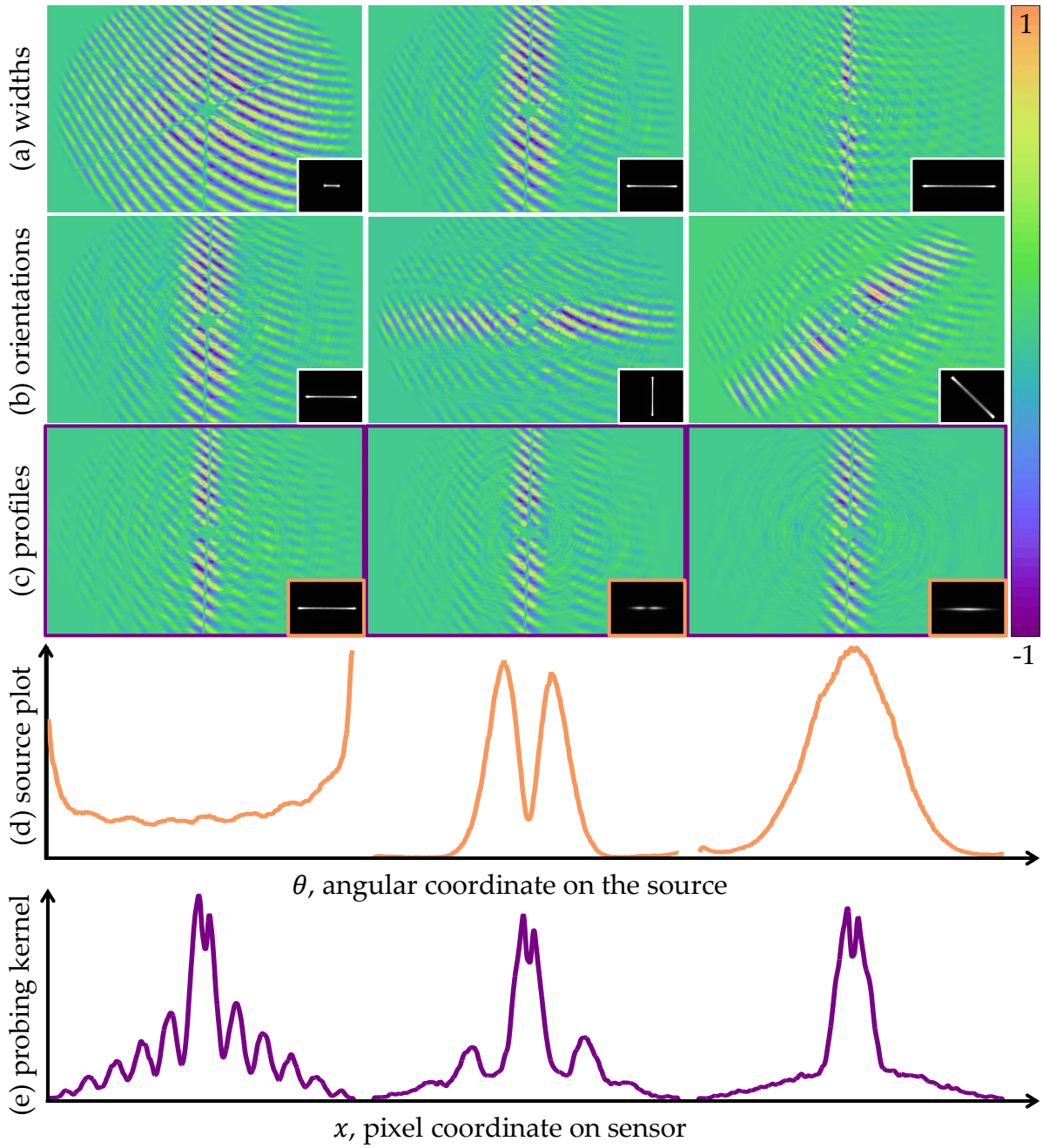


Figure 3.3: **Coherent probing of a retroreflector scene using linear source profiles.** The images show the real part of transmission matrices probed using the amplitude modulation functions $A(\theta)$ shown in the inset. Going horizontally, they vary respectively **(a)** source width, **(b)** source orientation and **(c)** source profile. **(d)** plots the source profiles in (c), and **(e)** plots the corresponding probing functions.

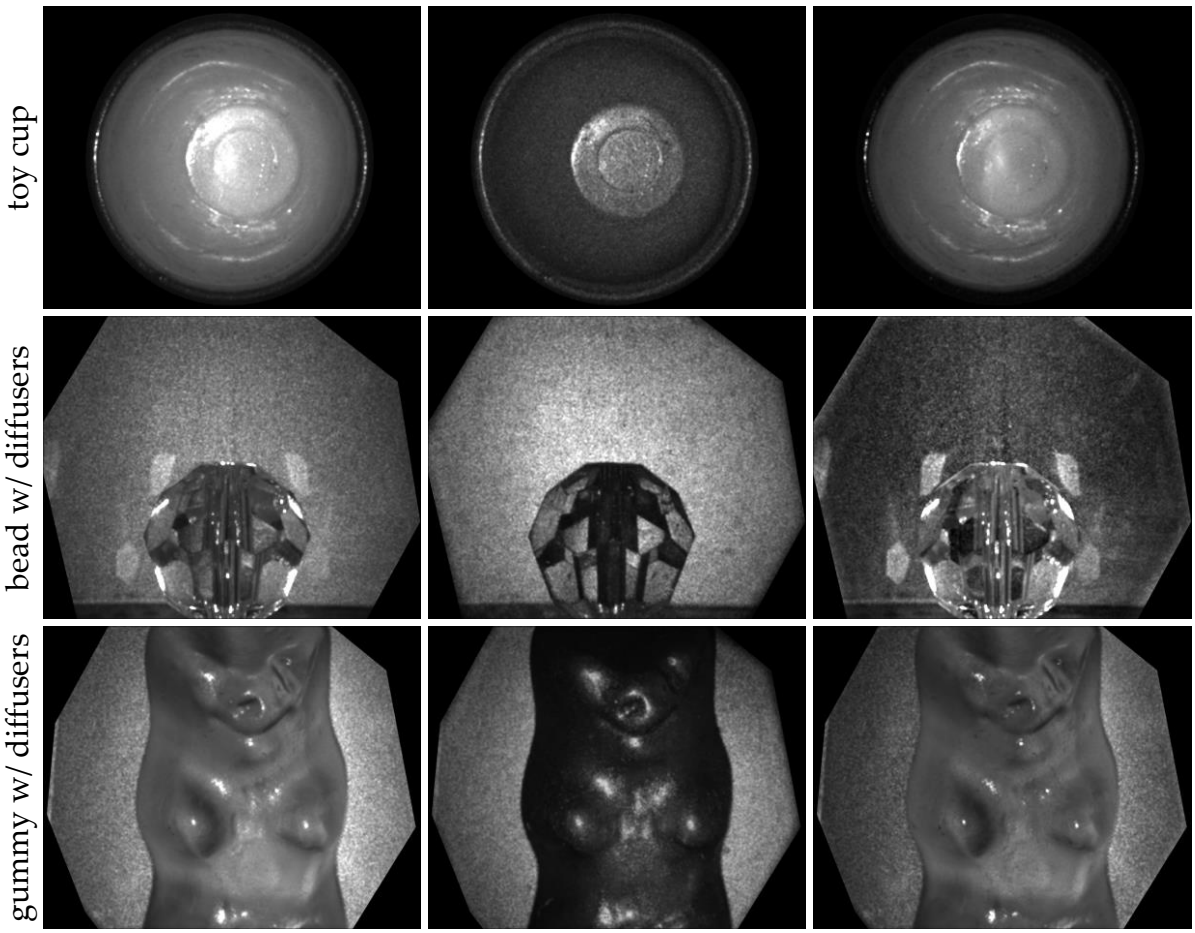
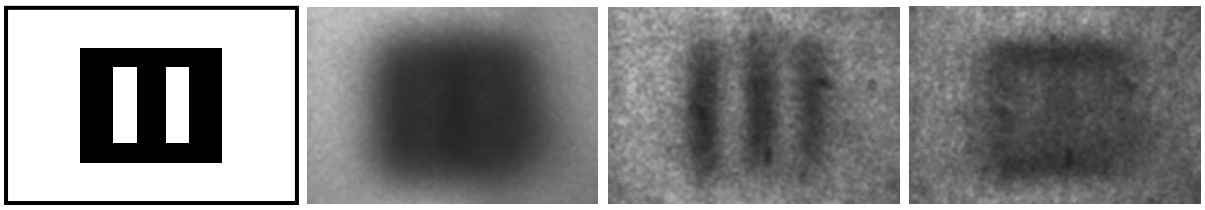


Figure 3.4: **Direct-indirect separation.** We use three scenes: cup (top), bead (middle), and gummy bear (bottom). The columns show **(a)** a regular photograph of the scene, **(b)** direct-only and **(c)** indirect-enhanced images captured using our probing technique. The cup and gummy bear appear metallic in the direct-only images, as the subsurface scattering effects dominating their appearance are only measured in the indirect-enhanced images. The direct-only image of the bead is missing the strong specular reflections on the bead's faces and the on the wall, which are prominent in the indirect-enhanced image.

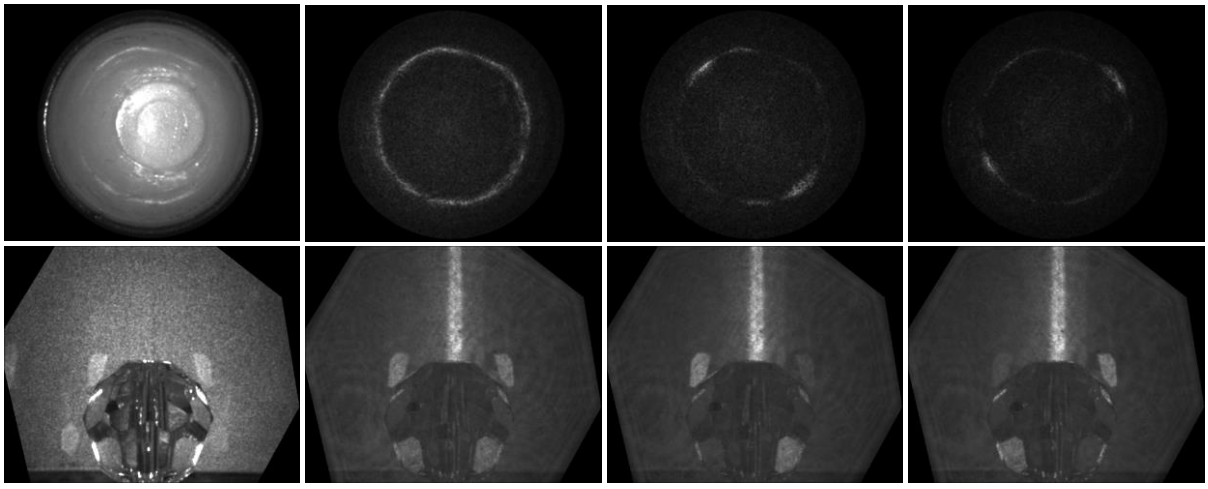


(a) pattern hidden behind scatterer | (b) scene-only camera image | (c) probing with horizontal line | (d) probing with vertical line

Figure 3.5: **Anisotropic descattering. (a)** We place a pattern consisting of two vertical stripes behind a thick semi-transparent scatterer. **(b)** In a regular image of the scene, the stripe features cannot be resolved because of the scattering. **(c)** Probing the scene with a horizontal Laplacian-of-Gaussian function that matches the orientation of the two stripes enhances their contrast and makes them clearly visible. **(d)** By contrast, probing the scene with a vertical function results in the two stripes remaining obscured.

Anisotropic descattering. As noted in O'Toole et al. [50], when imaging a target through a scatterer, it is possible to suppress the effect of scattering by probing with a pattern that emphasizes the diagonal of the light transport matrix, and subtracts the first few off-diagonals. Using coherent probing, we can perform this type of descattering with a probing function that is shaped like a Laplacian-of-Gaussian function – the positive central lobe emphasizes the main diagonal of the transmission matrix, and the off-center negative values subtract the off-diagonals. We can additionally selectively perform descattering along only certain directions by using probing functions that are shaped like *anisotropic* Laplacian-of-Gaussian functions. This is equivalent to a form of optical matched filtering that emphasizes features similar to the probing function.

Figure 3.5 shows an example of this optical matched filtering procedure. We use a template consisting of two vertical white stripes (width 100 µm) against a black background (Figure 3.5(a)). We place this template behind a semi-transparent scatterer of thickness 50 mm, created by immersing titanium-dioxide microspheres in silicone rubber. In a regular image of this scene (Figure 3.5(b)), the two stripe features are obscured by the scattering and are not discernible. Probing with a horizontal Laplacian-of-Gaussian function enhances vertical features, making it possible to resolve the stripe patterns (Figure 3.5(c)). By contrast, the stripes remain obscured when we probe with a vertical function (Figure 3.5(d)). We note that, in contrast to our technique, these anisotropic probing functions cannot be implemented in a light-efficient way using primal-dual coding.



(a) full scene (b) no probing (c) +45⁰ line source (d) -45⁰ line source

Figure 3.6: **Combining probing and transient imaging.** We use two scenes: cup (top), and bead (bottom). From left to right: regular image, and isolated frames of the transient sequence at the same time instance when using no probing, anisotropic probing at $+45^\circ$, and anisotropic probing at -45°. In the cup scene, when using probing, only the parts of the caustic that match the probing function are maintained, and the rest of the caustic is suppresed. Likewise, in the bead scene, different reflections on wall and on the bead faces are maintained and suppressed, depending on the probing function used. Please see the project website [34] for the complete transient sequences and additional probing patterns.

Combining probing and transient imaging. We perform experiments where we use temporally incoherent light to combine our probing technique with transient imaging. We use our system to capture transient sequences without any probing, as well as sequences combined with anisotropic probing at different orientations. Figure 3.6 shows examples for two scenes, where we isolate the same frame in three transient sequences captured with different probing patterns. We observe that, when using anisotropic probing, only features that match the probing function are preserved, while other features are suppressed relative to their intensity when no probing is used. We provide video visualizations of the entire transient sequences at the project website [34].

Measuring the entire light transport matrix. We create a two-dimensional scene consising of a diffuse surface and a mirror forming a right-angle corner (Figure 3.7(a)). For convenience, and for the limited resolution of the experiment, we will talk of transport matrices instead of functions. The two-dimensional light transport matrix of this scene is dominated by the main diagonal and anti-diagonal (Figure 3.7(c)): The diagonal of the

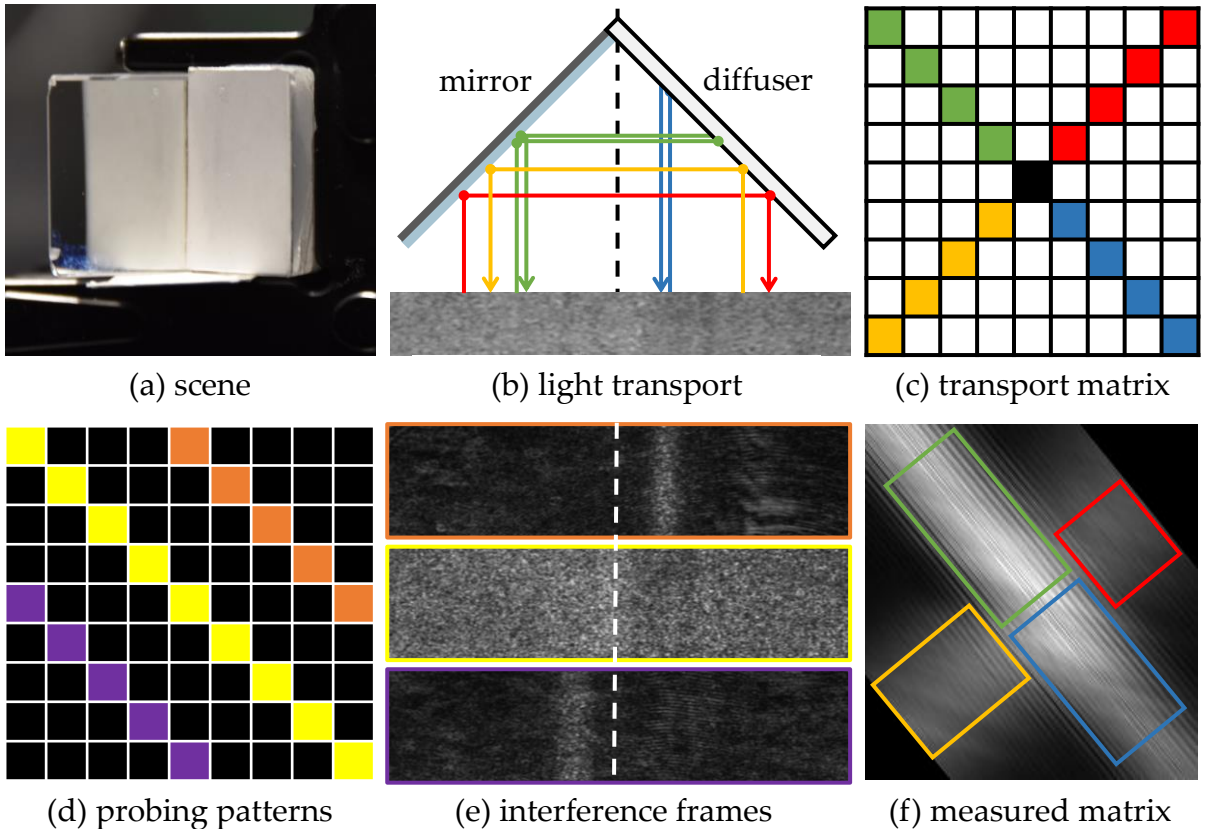
light transport matrix corresponds to contributions from direct and retroreflecting paths, whereas the anti-diagonal corresponds to contributions from two-bounce reflection paths (Figure 3.7(b)).

We use our technique to capture the two-dimensional light transport matrix of this scene (Figure 3.7(d)). We achieve this by successively probing with shifted diagonal functions (Figure 3.7(e)), implemented using phase modulation $\Phi(\theta)$ equal to ramp functions of varying slopes. The correlation measurements captured for different slopes (Figure 3.7(f)) are equal to different diagonals of the light transport matrix, and stacking them together forms the entire light transport matrix. We provide a video visualization of this procedure at the project website [34]. We note that the probing functions in this experiment correspond to the short-range-indirect probing patterns that have recently been successfully used for tissue imaging [19, 35, 40].

3.7 Discussion

Comparison with primal-dual coding. Interferometry with coded mutual intensity and primal-dual coding [50] both implement probing of incoherent light transport, albeit using very different optical configurations and operating principles. Therefore, it is worthwhile discussing their relative strengths and weaknesses. To be consistent with the terminology of O'Toole et al. [50], we will refer to the transmission, transport and probing *matrices*, which are spatially-quantized versions of the transmission, transport and probing functions.

We first contrast how the two methodologies implement different probing matrices. Given a desired probing matrix, primal-dual coding decomposes it into a sequence of illumination and sensing masks, which are then applied respectively on a projector source and sensor, successively within a single exposure. As discussed by O'Toole et al. [53], this procedure is light-efficient only when the probing matrix decomposition is *homogeneous*; for arbitrary probing matrices, a homogeneous decomposition may not exist or may require a very large number of mask patterns. On the other hand, interferometry with coded mutual intensity uses an inverse Fourier transform to convert to a probing matrix to individual amplitude and phase modulation functions, which are then applied to the Fourier-domain of the illumination and reference arm, respectively, of an interferometer. From a hardware complexity perspective, the two techniques introduce a trade-off between sensor masking (primal-dual coding) versus interferometry with phase modulation (interferometry with coded mutual intensity). In terms of the



(a) scene (b) light transport (c) transport matrix

(d) probing patterns (e) interference frames (f) measured matrix

Figure 3.7: **Measuring the light transport matrix of a mirror-diffuser corner scene. (a)** Photograph of the scene for visualization. **(b)** Light paths in the scene: blue are direct paths, red and yellow are two-bounce reflection paths, and green are retroreflecting paths. Below the schematic is a regular image of the scene. **(c)** The two-dimensional light transport matrix for this scene: direct and retroreflecting paths show up on the diagonal, and the two-bounce reflection paths show up on the antidiagonal. **(d)** Measured light transport matrix, with regions corresponding to different matrix parts in (c) marked in color. **(e)** Example probing patterns used to measure the light transport matrix. **(f)** Measurements using the probing patterns in (e), where the top and bottom images are due to two-bounce reflection paths, and the middle image is due to direct and retroreflecting paths. The dashed lines in (b) and (f) indicate the corner location. Please see the project website [34] for the complete set of probing measurements.

types of probing that can be realized, by using a Fourier-domain redistributive projector for amplitude modulation (Section 3.4), interferometry with coded mutual intensity enables light-efficient implementations of probing matrices that would be challenging to implement with primal-dual coding (for example, the anisotropic probing patterns shown in Section 4.4). However, this comes at the cost of only being able to implement Toeplitz probing matrices, a limitation not shared by primal-dual coding. Additionally, the advantage of better light efficiency can be outweighed by other noise performance issues typical in interferometry, as we discuss later in this section.

As probing the spatial dimensions of light transport is orthogonal to probing its temporal dimension, both primal-dual coding and interferometry with coded mutual intensity can be combined with transient imaging techniques. Primal-dual coding is better suited for macroscopic transient imaging, as the optical components it requires can be readily combined with continuous-wave time-of-flight cameras [1, 51]. By contrast, interferometry with coded mutual intensity is more readily combined with microscopic transient imaging, as both techniques rely on interferometric setups [18].

Finally, there are capabilities that have been demonstrated using only one or the other methodology: Primal-dual coding has been used in non-confocal setups to implement epipolar and disparity probing patterns [52]. By contrast, we have only used interferometry with coded mutual intensity in a confocal setting, though our theory can also be applied to non-confocal settings (for example, by incorporating our amplitude and phase modulation systems within a Mach-Zehnder interferometer). Conversely, we have shown that interferometry with coded mutual intensity can be used to perform coherent probing of transmission matrices. This capability has not been demonstrated using primal-dual coding, though O'Toole [48] discusses possible extensions to this end.

Noise and speckle. The signal-to-noise ratio of measurements captured using interferometry with coded mutual intensity can significantly deteriorate because of two factors inherent in interferometric systems. The first factor relates to the fact that the measurements captured by cameras in interferometric systems include not only the interference component, but additionally the DC intensity images of the reference and scattered arms (Equation (3.29)). The Poisson noise in these measurements is proportional to the total intensity [23], and thus it can potentially dominate the interference component when that is significantly weaker than the DC component. We refer to Takada [65] for a detailed analysis of the noise characteristics of interferometric measurements.

The second factor relates to the presence of speckle due to the use of (partially-

)coherent illumination. As Gkioulekas et al. [18] have shown, speckle can negatively impact interference contrast, and careful optical design is needed to minimize its effect. The impact speckle has on the quality of interferometric measurements can be assessed from the effect speckle has on an image of the target arm under the same illumination as that used for interferometry.

In practice, we note that both of these issues are counteracted to a significant extent by the fact that we estimate the phase and amplitude of the interference component using multiple image measurements, captured with sub-wavelength phase shifts (Section 3.5). However, as we discuss immediately below, this comes at the cost of slower operation of the imaging system.

Towards real-time operation. Our current implementation captures images at a relatively low frame rate of 6 Hz. The primary factor preventing us from achieving higher frame rates is our use of a phase SLM to implement not only phase modulation (which remains constant throughout capture), but also sub-wavelength phase shifts for estimating the interference component. Consequently, the maximum refresh rate of the phase SLM (60 Hz for commercially-available units) creates an upper bound on the maximum frame rate that can be achieved by our system.

To overcome this constraint, we can draw inspiration from other applications of phase-shifting interferometry. For example, we could potentially use the phase SLM to implement space-division multiplexing [28], allowing us to capture multiple phase shifts in a single measurement, at the cost of reduced spatial resolution. Alternatively, we could apply phase retrieval algorithms [16], in order to recover the phase and amplitude of the interference component from measurements at just one or two phase shifts. A potential downside of these approaches is that, by reducing the number of measurements, they can result in significant deterioration of signal-to-noise ratio, as discussed above.

A different approach would be to accelerate measurements at multiple phase-shifts, by replacing the phase SLM with a different optical component. In particular, we could potentially use a combination of a phase EOM with a MEMS mirror at the reference arm of our interferometer to implement phase modulation at kilohertz rates, analogously to our implementation of amplitude modulation.

3.8 Summary

We introduced a new imaging methodology, interferometry with coded mutual intensity, to perform both coherent probing of light transmission matrices, and incoherent probing of light transport matrices. Our theory and optical design allow us to programmatically implement arbitrary convolutional probing patterns, by applying appropriate amplitude and phase modulation at the focal plane of the illumination lens in a Michelson interferometer, without the need for hardware changes. These capabilities can additionally be incorporated into systems that use temporally-incoherent illumination for optical coherence tomography. We have developed a physical prototype, and demonstrated the validity of our theory and utility of our designs in experiments such as visualizing complex fields, capturing direct and global transport components, acquiring light transport matrices, and performing anisotropic descattering, both in steady-state and transient imaging. We hope that, by significantly expanding the types of probing that can be achieved using interferometric techniques, and by providing the design and engineering details for developing physical prototypes implementing these techniques, this work will enable further research on interferometric light transport techniques.

Chapter 4

Swept-angle synthetic wavelength interferometry

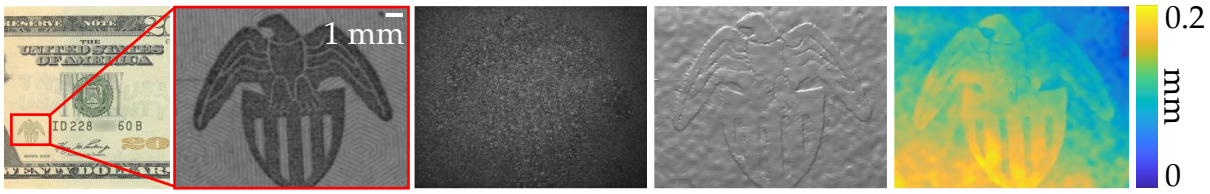


Figure 4.1: **Reconstructing the eagle embossed on a United States twenty-dollar bill**. The features on the eagle are raised 20 µm off the surface of the bill. From left to right: location of the eagle on the bill, an ambient light image of the eagle, a raw frame captured by our method, the depth recovered by our method rendered as a surface, and plotted as a depth map. The state-of-the-art technique, optical coherence tomography, requires twenty times as many measurements as our method does to capture a depth map of comparable quality.

We will now examine an application of joint spatio-temporal interferometric transmission probing to a core problem in computer vision: depth sensing. A variety of techniques is available for acquiring depth maps of three-dimensional scenes at different scales. Our focus here is upon centimeter-sized scenes with features that need to be imaged at micrometer-scale resolutions. Existing passive (stereo, defocus, shading) and active (lidar, structured light, and correlation time-of-flight (ToF)) illumination depth sensing techniques cannot provide micrometer axial resolution: Limitations in electronics technology and imaging noise constrain these to hundreds of micrometers of accuracy. Interferometry, applicable naturally at micrometer scales, is therefore the ideal choice for

micron-scale depth sensing.

As we saw in Section 2.3, the choice of illumination spectrum in the interferometer leads to techniques such as optical coherence tomography (OCT), which uses broadband illumination, and phase-shifting interferometry (PSI), which uses monochromatic illumination. On one extreme, PSI achieves nanometer-scale resolution but provides unambiguous depth measurements only for scenes less than one wavelength in depth range. On the other, OCT achieves micrometer-scale resolution over arbitrary depth ranges. We consider synthetic wavelength interferometry (SWI), which operates between these two extremes: By using illumination consisting of two narrowly-separated optical wavelengths, SWI provides a controllable trade-off between the large unambiguous depth range of OCT, and the large axial resolution of PSI. SWI can achieve micrometer resolution at depth ranges in the order of hundreds of micrometers.

A factor that complicates depth recovery from interferometry is the presence of indirect light that scatters multiply in the scene. As we will see in Figure 4.4, indirect light results in depth recovery errors. Traditional implementations of these techniques therefore use beam-scanning and focusing on scene points one at a time, which slows down acquisition and therefore decreases either the lateral extent or the lateral resolution of the recovered depth. On the other hand, full-field techniques capture the entire lateral extent of the scene in one shot, but are strongly affected by indirect light.

We introduce *swept-angle synthetic wavelength interferometry*, a spatially-probed version of SWI, that introduces robustness to indirect light into interferometric measurements. As summarized in Table 4.1, swept-angle SWI achieves fast depth sensing robust to indirect light, with unambiguous depth range tunable to our requirement.

Table 4.1: **Comparison of interferometric depth sensing techniques for millimeter-size scenes**.

method	unambiguous depth range	acquisition time	robustness to indirect light
full-field TD-OCT	✓	✗	✓
scanning FD-OCT	✓	✗	✓
scanning SS-OCT	✓	✗	✓
full-field PSI	✗	✓	✗
scanning PSI	✗	✗	✓
full-field SWI	✓	✓	✗
scanning SWI	✓	✗	✓
full-field SA-SWI	✓	✓	✓

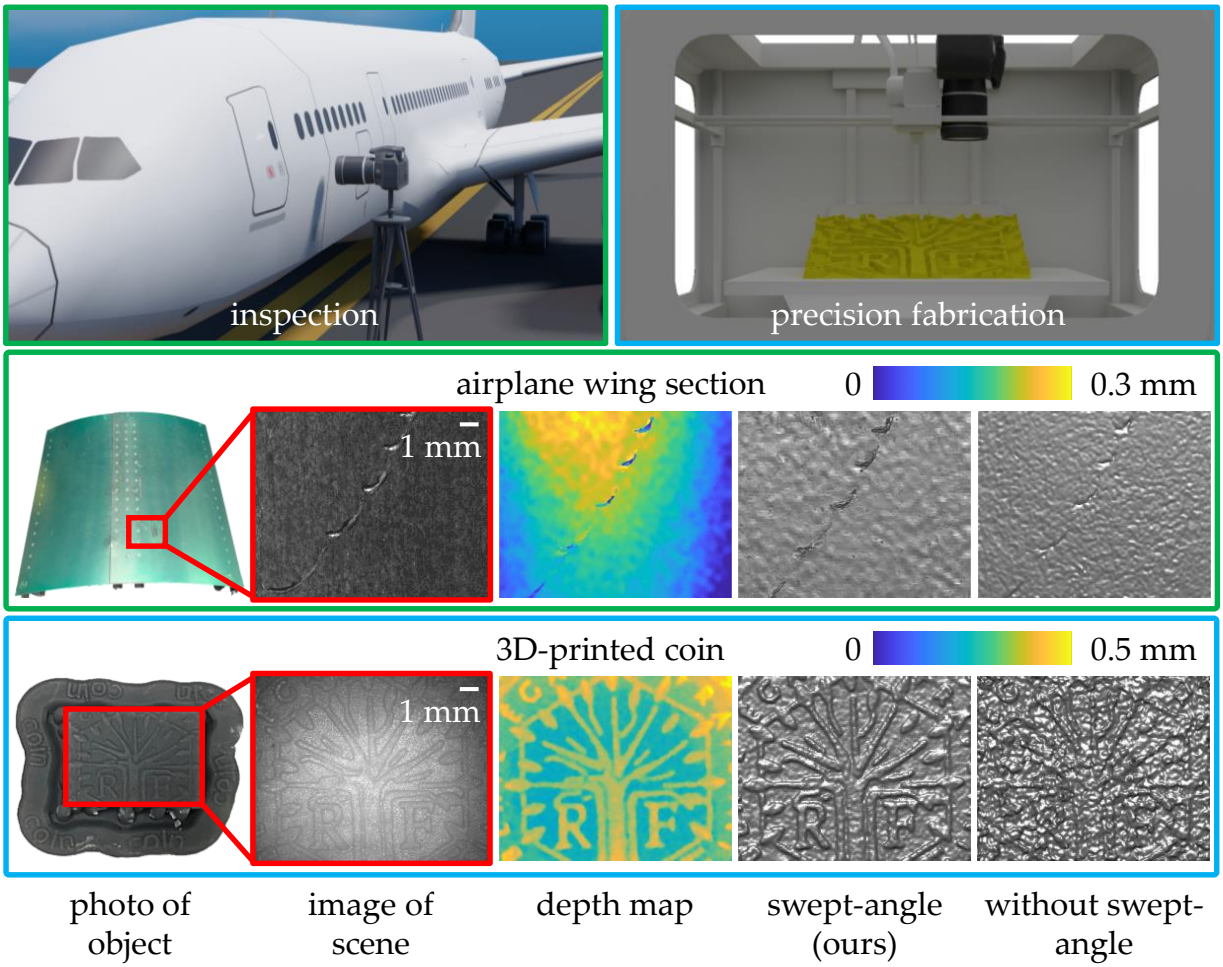


Figure 4.2: **Applications of swept-angle SWI in industrial inspection and fabrication.** We show depth reconstructions for two scenes representative of these applications: millimeter-scale dents on an airplane wing section, and a 3D-printed coin.

Potential applications. Swept-angle SWI can be useful and relevant for critical applications, including industrial fabrication and inspection. In industrial fabrication, swept-angle SWI can be used to provide feedback during additive and subtractive manufacturing processes [63]. In industrial inspection, swept-angle SWI can be used to examine newly-fabricated or used in-place critical parts and ensure they comply with operational specifications. Swept-angle SWI, uniquely among 3D scanning techniques, offers a combination of operational characteristics that are critical for both applications: First, high acquisition speed, which is necessary to avoid slowing down the manufacturing process, and to perform inspection efficiently. Second, micrometer lateral and axial resolution, which is necessary to detect critical defects. Third, robustness to aberrations and indirect illumination, which is necessary because of the materials used for

fabrication, which often have strong subsurface scattering. Figure 4.2 showcases a result representative of these applications: We use swept-angle SWI to scan a wing section from a decommissioned Boeing 747, to detect critical defects such as scratches and bumps from collisions, at axial and lateral scales of a couple dozen micrometers. We also use swept-angle SWI to scan a coin pattern 3D-printed by a commercial material printer on a translucent material.

4.1 Synthetic wavelength interferometry

Synthetic wavelength interferometry uses illumination comprising two distinct, but narrowly-separated, optical wavelengths. We denote these wavelengths as λ and $\lambda/1+f$, corresponding to wavenumbers $\kappa \equiv 2\pi/\lambda$ and $(1+f)\,\kappa$, respectively, where $f \ll 1$. As in a standard full-field Michelson interferometer, we assume that this illumination is injected into the interferometer as a collimated beam – for example, created by placing the outputs of two fiber-coupled single-frequency lasers at the focal plane of a lens, as in Figures 1.2 and 4.3. This illumination has the spectrum

$$S\left(\tilde{\kappa}\right) = \delta\left(\tilde{\kappa} - \kappa\right) + \delta\left(\tilde{\kappa} - \kappa(1+f)\right), \tag{4.1}$$

where δ is the Dirac delta function. As before, we will derive the coherence characteristics of the illumination. Since all light per wavelength arises from a single point, this illumination is fully spatially coherent: $\Gamma\left(\epsilon\right) = 1$. From Equation (2.15), the temporal coherence function of this illumniation is the Fourier transform of the spectrum:

$$\mathcal{G}\left(s\right) = \exp\left(-\mathrm{j}\kappa s\right) + \exp\left(-\mathrm{j}\kappa(1+f)s\right). \tag{4.2}$$

For the moment, we will assume that the scene has a transmission function

$$\mathcal{T}^c\left(x, x', \tau\right) = \delta\left(x - x'\right)\delta\left(\tau - 2d(x)/c\right). \tag{4.3}$$

$d\left(\tilde{x}\right)$ is defined as the minimum z any point on the scene intersects with the line $x = \tilde{x}$. We will discuss the implications of this assumption later. With this transmission function,

Equation (2.18) gives us

$$\begin{aligned}\mathcal{C}(x) &= \int_{\epsilon,\tau} \delta(x - x')\,\delta(\tau - {}^{2d(x)}/_{c})\,[\exp(-\mathrm{j}\kappa(c\tau - 2l)) \\ &\qquad + \exp(-\mathrm{j}\kappa(1+f)(c\tau - 2l))]\,\mathrm{d}x'\,\mathrm{d}\tau \\ &= \exp(-\mathrm{j}\kappa(2d(x) - 2l)) + \exp(-\mathrm{j}\kappa(1+f)(2d(x) - 2l)) \\ &= \exp(-2\mathrm{j}\kappa(d(x) - l)) + \exp(-2\mathrm{j}\kappa(1+f)(d(x) - l)) \end{aligned} \tag{4.4}$$

The interference component at the camera is the real part of the correlation:

$$\begin{aligned}\mathrm{Re}\{\mathcal{C}(x)\} &= \mathrm{Re}\{\exp(-2\mathrm{j}\kappa(d(x) - l)) + \exp(-2\mathrm{j}\kappa(1+f)(d(x) - l))\} \\ &= \cos(2\kappa(d(x) - l)) + \cos(2\kappa(1+f)(d(x) - l)) \\ &= 2\cos(\kappa(2+f)(d(x) - l))\cos(\kappa f(d(x) - l)) \\ &\approx 2\cos(2\kappa(d(x) - l))\cos(\kappa f(d(x) - l)). \end{aligned} \tag{4.5}$$

We observe that, as a function of $d(x) - l$, the interference is the product of two sinusoids: first, a *carrier sinusoid* with *carrier wavelength* $\lambda_c \equiv {}^{\lambda}/_{2}$ and corresponding *carrier wavenumber* $\kappa_c \equiv 2\lambda$; second, an *envelope sinusoid* $\boldsymbol{\mathcal{E}}$ with *synthetic wavelength* $\lambda_s \equiv {}^{\lambda}/_{f}$ and corresponding *synthetic wavenumber* $\kappa_s \equiv \kappa f$:

$$\boldsymbol{\mathcal{E}}(x, l) \equiv \cos(\kappa f(d(x) - l)). \tag{4.6}$$

From Equation (4.6), we see that SWI encodes scene depth $d(x)$ in the phase $\phi(d(x)) \equiv \kappa f(d(x) - l)$ of the envelope sinusoid. We defer details on how to measure the envelope and estimate this phase until Section 4.3. We make two observations: First, SWI provides depth measurements at intervals of $[0, \lambda_s]$, and cannot disambiguate between depths differing by an integer multiple of λ_s. Second, the use of two wavelengths makes it possible to control the unambiguous depth range: By decreasing the separation κf between the two emitted wavelengths, we increase the unambiguous depth range, at the cost of decreasing depth resolution.

Implications of Equation (4.3). The transmission function in Equation (4.3) assumes two conditions being met. Firstly, the term $\delta(x - x')$ implies that only light originating at x contributes to scene appearance. Secondly, the term $\delta(\tau - {}^{2d(x)}/_{c})$ implies that all of the scene's appearance is reflected off of the layer of the scene closest to the beamsplitter. In the ideal case where both of these are fulfilled, synthetic wavelength interferometry

can estimate accurate $d\,(x)$. However, these conditions are easy to violate: For instance, in highly scattering scenes, the presence of subsurface scattering can cause the first assumption to break down. In addition, imaging depth through slightly reflective transparent layers can cause the second assumption to break down. As we will see in the next section, some interferometers are designed to physically reject light that violates these assumptions: However, their designs cause significant limitations in the quality of the acquired depth.

4.2 SWI implementations for depth sensing

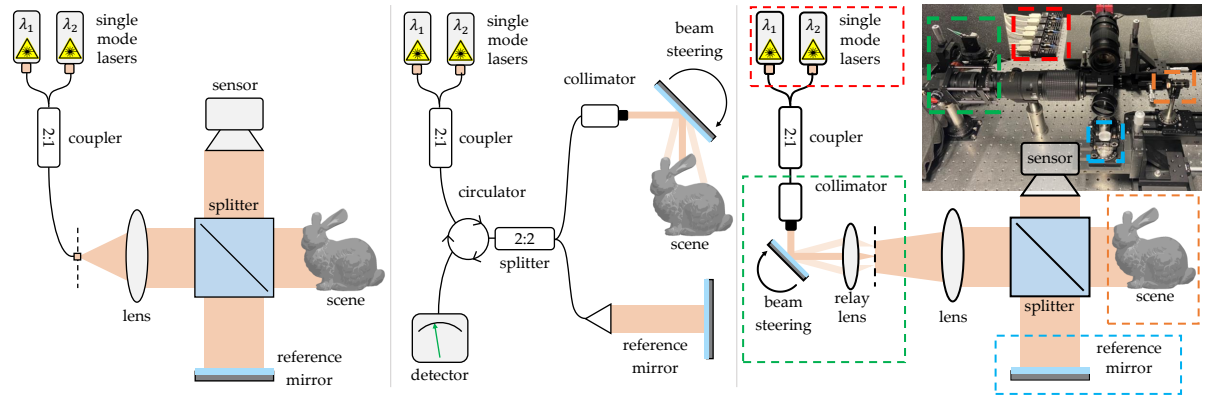
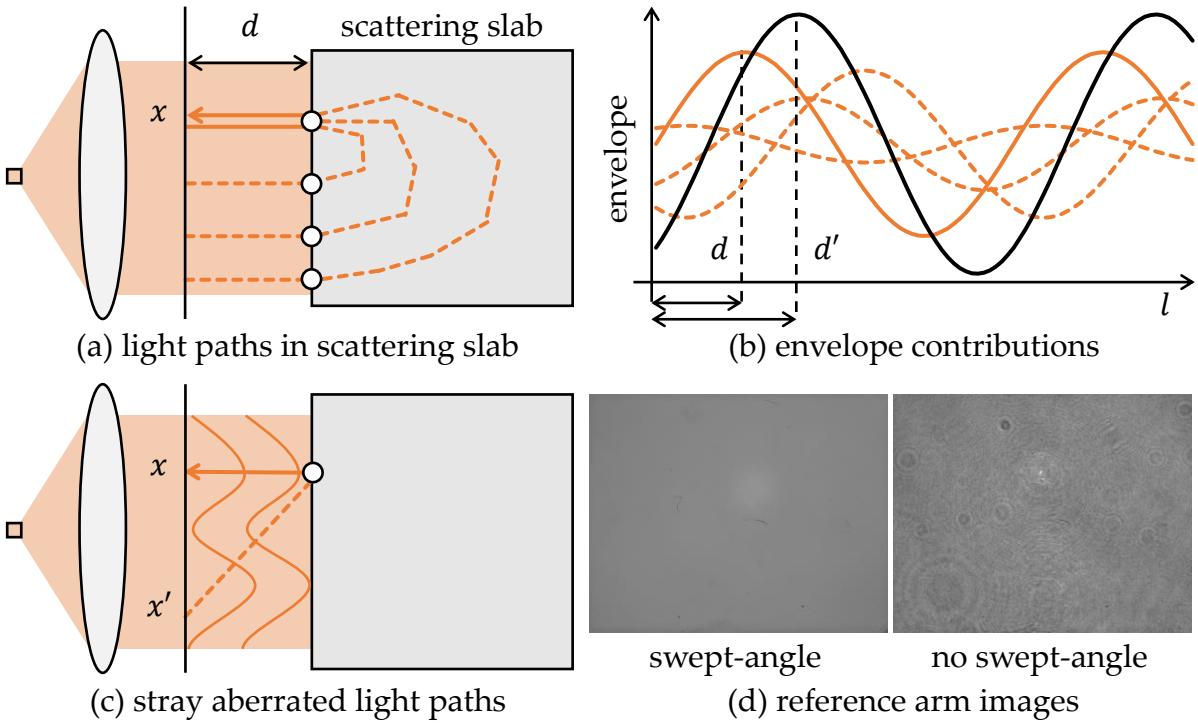


Figure 4.3: **Synthetic wavelength interferometry setups. (a)** A full-field interferometer efficiently acquires full-frame depth, but is susceptible to phase corruptions due to aberrations and indirect illumination. **(b)** A scanning interferometer is robust to such corruptions, but requires slow lateral scanning. **(c)** A swept-angle full-field interferometer achieves both efficiency and robustness.

Figure 4.3 shows two types of Michelson interferometer setups that implement SWI: (a) a full-field interferometer; and (b) a scanning interferometer. We discuss their relative merits, which will motivate our spatial probing-based swept-angle interferometer setup.

Full-field SWI. As we saw before, full-field interferometers (Figure 4.3(a)) use *free-space optics* (lenses, beamsplitters) to illuminate and image the entire field of view in the scene and reference arms. They also use a two-dimensional sensor to measure interference at all locations x. This enables fast interference measurements for all scene points at once, and at lateral resolutions as high as the pixel pitch of the sensor.

Unfortunately, full-field interferometers are susceptible to phase corruption effects, as we visualize in Figure 4.4. Equation (4.3) assumes that the scene field is due to only



(a) light paths in scattering slab (b) envelope contributions

(c) stray aberrated light paths (d) reference arm images

Figure 4.4: **Phase corruption effects. (a)** A typical scene contains both direct (solid line) and indirect (dashed lines) light paths. **(b)** The direct path contributes a sinusoid with the correct phase (solid orange line). The indirect paths contribute sinusoids with incorrect phases (dashed lines). Their summation results in erroneous phase estimation (dark solid line). **(c)** Aberrations in free-space optics result in stray light paths that also contribute incorrect phases. **(d)** Images of the reference arm (mirror) visualize the effects of aberrations, and their mitigation using swept-angle illumination.

the *direct* light path (solid line in Figure 4.4(a)), which produces a sinusoidal envelope with phase delay $d\,(x)$ (solid orange curve in Figure 4.4(b)). In practice, the scene field will include contributions from additional paths: First, *indirect* paths due to subsurface scattering (dashed lines in Figure 4.4(a)). Second, stray paths due to optical aberrations (Figure 4.4(c)). These paths have different lengths, and thus contribute to the envelope sinusoidal terms of different phases (dashed curves in Figure 4.4(b)). Their summation produces an overall sinusoidal envelope (black) with phase $d' \neq d\,(x)$, resulting in incorrect depth estimation.

Scanning-based SWI Scanning interferometers (Figure 4.3(b)) use *fiber optics* (couplers, circulators, collimators) to generate a focused beam that illuminates only one point in the scene and reference arms. Additionally, they use a single-pixel sensor, focused at the same point. They also use steering optics (for example, MEMS mirrors) to scan

the focus point and capture interference measurements for the entire scene. Scanning interferometers effectively mitigate the phase corruption effects in Figure 4.4: Because, at any given time, they only illuminate and image one point in the scene, they eliminate contributions from indirect paths such as those in Figure 4.4(a). Additionally, because the use of fiber optics minimizes aberrations, they eliminate stray paths such as those in Figure 4.4(c). Unfortunately, this robustness comes at the cost of having to use beam steering to scan the entire scene. This translates into long acquisition times, especially when it is necessary to measure depth at pixel-level lateral resolutions and at a sensor-equivalent field of view. This scan is challenging for a variety of reasons, a few of which we will discuss here.

Performing a micrometer lateral scan requires: (1) a laser beam that can be collimated or focused to a spot size of a few micrometers; (2) a MEMS mirror capable of scanning at high-enough angular resolution to translate the laser beam a few microns on the scene surface; and (3) acquisition time long enough to scan a megapixel-size grid on the scene. Each of these requirements is difficult to achieve:

1. The diameter of a Gaussian laser beam is inversely proportional to its divergence [64, Chapter 4]. The smaller the beam diameter, the larger the divergence. At 780 nm, a laser beam with a diameter of 1 μm grows in diameter by 10% every 2 m. Therefore, maintaining collimation diameter of 1 μm is challenging except for very small working distances.

   As an alternative to using a thin, collimated laser beams, we can use a beam that is focused at each point on the scene. Contrary to micron-scale beam waists, it is possible to focus single-mode lasers to pixel-size spot sizes [64, Chapter 9]. However, focusing the laser beam onto the scanned scene points sharply decreases the depth of field of the imaging system: Whereas, in the case of a collimated beam, the depth of field is limited by the divergence of the collimated beam, in the case of a focused beam, it is limited by the quadratic phase profile of the focused spot. Effectively, to use this focused setup, we need another axial scan to ensure that the scanned post is within the depth of field, which only adds to acquisition time.

2. Top-of-the-line scanning micromirrors typically have angular scanning resolutions of 10 μrad [46]. The maximum working distance such that this would correspond to micrometer lateral resolution is 10 cm.

3. The scanning micromirror needs to be run in "point-to-point scanning mode" [46], where the micromirror stops at every desired position. The best settling times for step mirror deflections are around 100 μs [46]. Using these numbers, for a

megapixel image, the micromirror rotations, excluding exposure time, require a total time around 100 s.

Full-field swept-angle SWI. We will now introduce full-field swept-angle SWI, which does not need to perform lateral scanning and is robust to indirect light by construction. The design of this system is based on the swept-angle mechanism introduced in Chapter 3.

The robustness of scanning interferometry to indirect light comes from its ability to physically reject indirect illumination from the measured scene appearance by illuminating and imaging only one point at a time. An alternative to rejecting from the sensor measurement itself is to reject it only from the interference component, an operation we can perform using the direct-only probing operation we set forth in Chapter 3. To see how this can be done, we start from Equation (3.3):

$$\mathcal{C}\left(x;\mathcal{P}^c\right) = \frac{1}{c}\int_{\epsilon,s} \mathcal{T}^c\left(x, x+\epsilon, {}^{s+2l}/_c\right)\mathcal{P}^c\left(\epsilon, s\right)\,\mathrm{d}\epsilon\,\mathrm{d}s. \tag{4.7}$$

If we were to set the probing function to be $\mathcal{P}^c\left(\epsilon, s\right) = \mathcal{G}\left(s\right)\delta\left(\epsilon\right)$, where δ is the Dirac delta function, the correlation is:

$$\begin{aligned}\mathcal{C}\left(x;\mathcal{P}^c\right) &= \frac{1}{c}\int_{\epsilon,s} \mathcal{T}^c\left(x, x+\epsilon, {}^{s+2l}/_c\right)\mathcal{G}\left(s\right)\delta\left(\epsilon\right)\,\mathrm{d}\epsilon\,\mathrm{d}s \\ &= \frac{1}{c}\int_{s} \mathcal{T}^c\left(x, x, {}^{s+2l}/_c\right)\mathcal{G}\left(s\right)\,\mathrm{d}s.\end{aligned} \tag{4.8}$$

This integral rejects any contributions from $\mathcal{T}^c\left(x, x', \tau\right)$ for $x \neq x'$, granting us the ability to perform imaging insensitive to indirect light and avoiding one of the assumptions in writing Equation 4.3. Unfortunately, the second assumption – that all (even direct-only) light comes from reflections off the closest surface of the scene – still needs to be made. However, as we will see in our results, this is a valid assumption for most real-world scenes, even those that have highly complex and scattering light transmission.

We then need to implement a spatial probing pattern $\Gamma\left(\epsilon\right) = \delta\left(\epsilon\right)$. This is easily done using the interferometer design in Section 3.3: We use Equation (3.25) to determine the required area emitter. The inverse Fourier transform of a Dirac delta is a constant function, which tells us we need a large, constant area emitter. In practice, we can create only a finite-sized emitter, and the theory in Section 3.3 is valid for small angles: Therefore, we can never implement an exact Dirac delta probing function. In our experiments, we use a square $1^\circ \times 1^\circ$ angular pattern. The Fourier transform of this pattern is a sinc, and

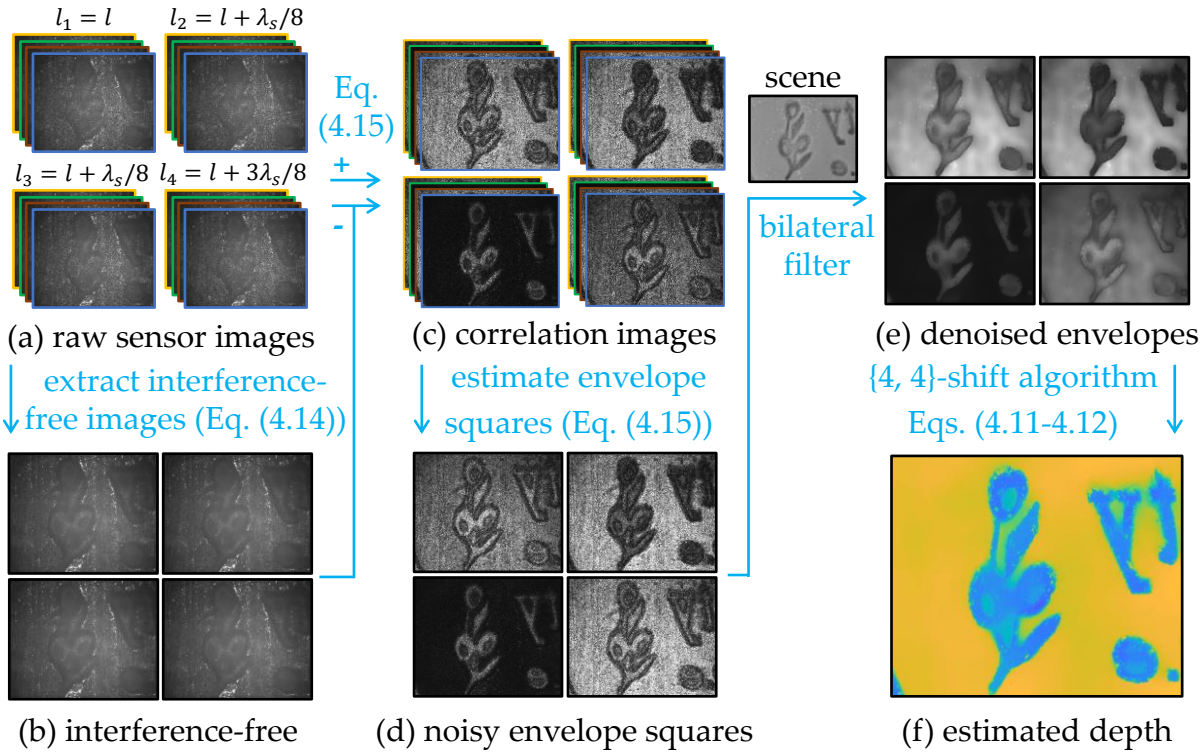


Figure 4.5: **The $\{4, 4\}$ phase retrieval pipeline. (a)** We take intensity measurements at 16 reference positions, corresponding to 4 synthetic by 4 carrier subwavelength shifts. **(b)** For each synthetic sub-wavelength shift, we estimate an interference-free image. **(c)** We subtract the interference-free image from sensor measurements to extract interference. **(d)** We extract the envelope from the interference images by estimating their squared magnitude. **(e)** We denoise the envelope image using joint bilateral filtering. **(e)** We use 4, 4-shift phase retrieval to estimate envelope phase and therefore depth.

the width of the sinc gives the effectivity of the indirect light rejection. We found in experiments that this source is enough for high-quality depth reconstruction.

4.3 Acquisition and depth recovery pipeline

We now discuss our pipeline for acquiring and processing image measurements to estimate the envelope phase, and thus depth. For this, in the setup of Figure 4.3(c), we will use a nanometer-accuracy translation stage to vary the location l of the reference mirror.

Phase retrieval. As we will see below, our acquisition pipeline has the capability to estimate the absolute squared magnitude of the envelope $|\mathcal{E}(x, l)|^2$. Using Equation (4.6),

we can express this as:

$$|\boldsymbol{\mathcal{E}}(x,l)|^2 = \cos^2(\kappa f(d(x)-l)) \tag{4.9}$$

$$= \frac{1+\cos(2\kappa f(d(x)-l))}{2}. \tag{4.10}$$

The period of this sinusoid is half the synthetic wavelength λ_s. Taking measurements at $l_n = l + {}^{n\lambda_s}/_{2N}$ for $N \in \{0,\dots,N-1\}$, we estimate the envelope phase using the classical *N-shift phase retrieval algorithm* [12, 36]:

$$\phi(d(x)) = \arctan\left[\frac{\sum_{n=0}^{N-1}|\boldsymbol{\mathcal{E}}(x,l_n)|^2 \sin({}^{2\pi n}/_N)}{\sum_{n=0}^{N-1}|\boldsymbol{\mathcal{E}}(x,l_n)|^2 \cos({}^{2\pi n}/_N)}\right]. \tag{4.11}$$

and the depth (up to an integer multiple of λ_s) as

$$d(x) = l + \frac{1}{2\kappa f}\arctan\left[\frac{\sum_{n=0}^{N-1}|\boldsymbol{\mathcal{E}}(x,l_n)|^2 \sin({}^{2\pi n}/_N)}{\sum_{n=0}^{N-1}|\boldsymbol{\mathcal{E}}(x,l_n)|^2 \cos({}^{2\pi n}/_N)}\right]. \tag{4.12}$$

Envelope estimation. Equation (4.11) requires estimates of the envelope $\boldsymbol{\mathcal{E}}(x,l_n)$ at each location l_n. To estimate these values, for each l_n, we capture intensity measurements $I(x,l_n^m)$ at M reference locations corresponding to shifts by M-th fractions of the *carrier* wavelength, $l_n^m = l_n + {}^{m\lambda_c}/_M$, $m \in \{0,\dots,M-1\}$. As the carrier wavelength λ_c is half the optical wavelength λ and orders of magnitude smaller than the synthetic wavelength λ_s, the envelope and interference-free image remain approximately constant across these shifts, $\boldsymbol{\mathcal{E}}(x,l_n^m) \approx \boldsymbol{\mathcal{E}}(x,l_n)$, $\mathbf{i}(x,l_n^m) \approx \mathbf{i}(x,l_n)$. Then, from Equations (1.5) and (4.5),

$$I(x,l_n^m) = \mathbf{i}(x,l_n) + 2\sin(\kappa_c(d(x)-l_n^m))\,\boldsymbol{\mathcal{E}}(x,l_n). \tag{4.13}$$

We estimate the interference-free image and squared magnitude of the envelope as

$$\mathbf{i}(x,l_n) = {}^1/_M \sum_{m=0}^{M-1} I(x,l_n^m), \tag{4.14}$$

$$|\boldsymbol{\mathcal{E}}(x,l_n)|^2 = {}^1/_{2M} \sum_{m=0}^{M-1} (I(x,l_n^m) - \mathbf{i}(x,l_n))^2. \tag{4.15}$$

Equations (4.14), (4.15), (4.11), and (4.12) conclude our phase estimation pipeline, which we summarize in Figure 4.5. We call this pipeline the $\{M,N\}$*-shift phase retrieval algorithm*. The parameters M and N are design parameters that we can fine-tune to control a trade-off between acquisition time and depth accuracy: As we increase M and N, the number $M \cdot N$ of total images we need to capture also increases; at the same

time, the final depth estimate becomes more robust to noise. The theoretical minimum number of images is achieved using $\{3, 3\}$ shifts, corresponding to 9 images. In practice, we found that $\{4, 4\}$ shifts perform robustly across all scenes we scanned. We evaluate different $\{M, N\}$ combinations in the experiments.

Dealing with speckle. Interference in non-specular scenes takes the form of *speckle*, a high-frequency pseudo-random pattern (Figure 4.5(a)). This can result in very noisy envelope, and thus phase and depth, estimates. As we show in Figure 4.4(d), our use of swept-angle illumination greatly mitigates the effects of speckle. We can further reduce the impact of speckle by denoising the estimated quantities with a low-pass filter (for example, Gaussian), as in prior work [18, 33]. Alternatively, to avoid blurring image details, we can use joint bilateral filtering [14, 56] with a guide image of the scanned scene captured under ambient light. Empirically, we found it advantageous to blur the envelope estimates of Equation (4.15) before using them in Equation (4.11) (Figure 4.5(d)), compared to blurring the phase or depth estimates.

4.4 Experiments

Experimental prototype. For all our experiments, we use the experimental prototype in Figure 4.3(c). Our prototype comprises: two distributed-Bragg-reflector lasers (wavelengths 780 nm and 781 nm, power 45 mW, linewidth 1 MHz); a galvo mirror pair; a translation stage (resolution 10 nm); two compound macro lenses (focal length 200 mm); a CCD sensor (pixel pitch 3.7 µm, 3400 × 2700 pixels); and other free-space optics and mounts. We provide a detailed component list and specifications, calibration instructions, and other information about the setup in the appendix.

We use the following experimental specifications: The reproduction ratio is 1:1, the field of view is 12.5 mm × 10 mm, and the working distance is 400 mm. The unambiguous depth range is approximately 500 µm. We use $\{4, 4\}$-shifts (16 images), and a minimum per-image exposure time of 10 ms, resulting at a frame rate of 5 Hz.

Depth recovery on challenging scenes. In Figures 4.6, 4.1 and 4.8, we show depth scans of a variety of challenging scenes. We scan materials ranging from rough metallic (coins), to diffuse (music box, pill), to highly-scattering (soap and chocolate). Most of our scenes include fine features requiring high lateral and axial resolution (music box, business card, US quarter, dollar bill). We compare scan results using swept-angle SWI

(with bilateral or Gaussian filtering), versus conventional full-field SWI (Figure 4.3(a), emulated by deactivating the angle-sweep galvo). In all scenes, the use of swept-angle illumination greatly improves reconstruction quality. The difference is more pronounced in scenes with strong subsurface scattering (music box, chocolate, soap). Even in metallic scenes where there is no subsurface scattering (coins), the use of swept-angle illumination still improves reconstruction quality, by helping mitigate aberration artifacts. Between bilateral versus Gaussian filtering, bilateral filtering helps preserve higher lateral detail.

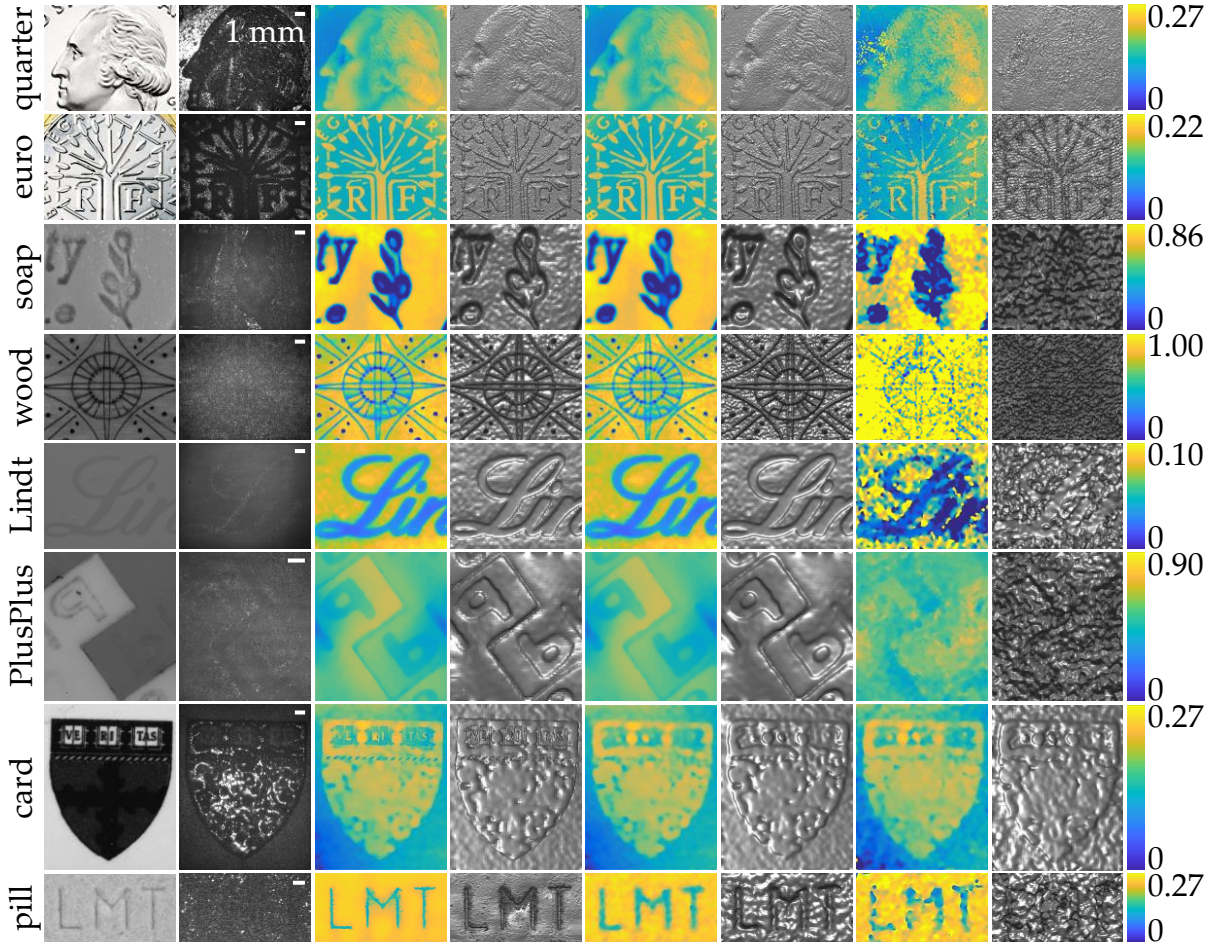


Figure 4.6: **Depth reconstruction on microscopic scenes**. Depth maps (left) and surface renderings (right) acquired using full-field SWI with: **(c)** swept-angle scanning and bilateral filtering; **(d)** swept-angle scanning and Gaussian filtering; **(e)** no swept-angle scanning and with Gaussian filtering. These are *microscopic* scenes, with millimeter-scale depth ranges. The colorbars are in millimeters.

In addition, we demonstrate that swept-angle SWI approximates the reconstruction

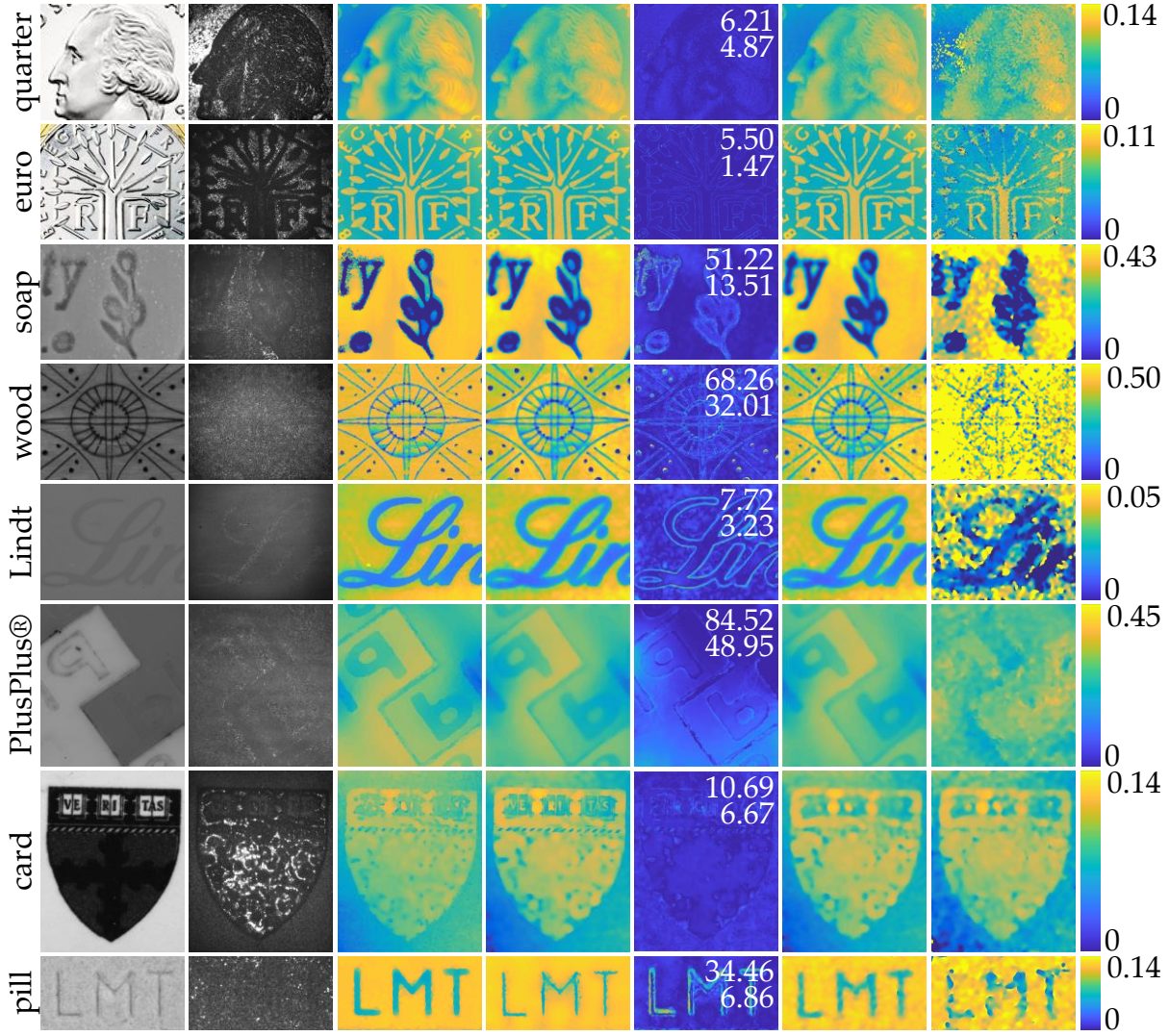


Figure 4.7: **Comparing swept-angle SWI and full-field OCT in microscopic scenes** (from Figure 4.6). The difference maps show the absolute difference between recovered OCT and swept-angle SWI depths, and report the root-mean-square (top) and mean absolute (bottom) differences between the two. The OCT depths are captured at a resolution comparable to swept-angle SWI, and should not be considered ground truth. The colorbars are in millimeters.

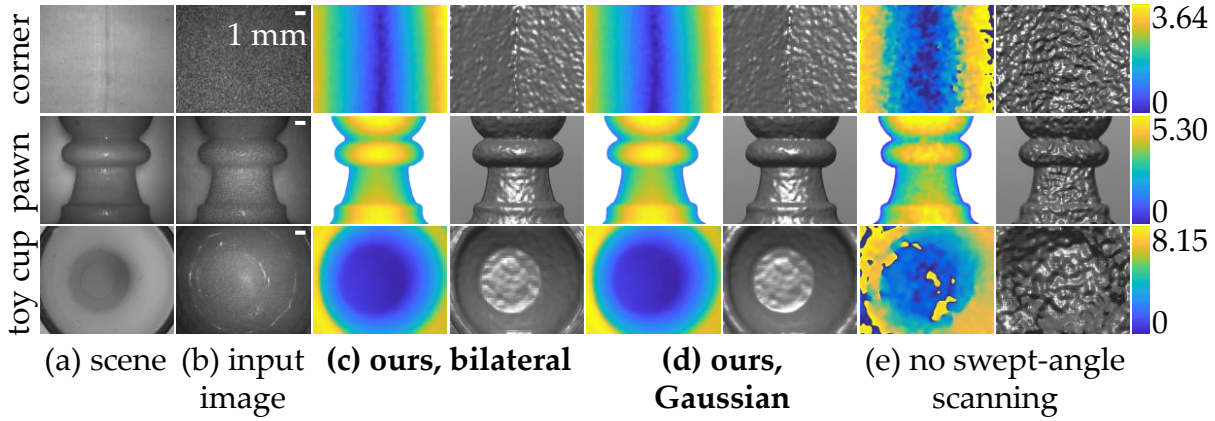


Figure 4.8: **Depth reconstruction on macroscopic scenes**. Depth maps (left) and surface renderings (right) acquired using full-field SWI with: **(c)** swept-angle scanning and bilateral filtering; **(d)** swept-angle scanning and Gaussian filtering; **(e)** no swept-angle scanning and with Gaussian filtering. These are *macroscopic* scenes, with centimeter-scale depth ranges. The colorbars are in millimeters.

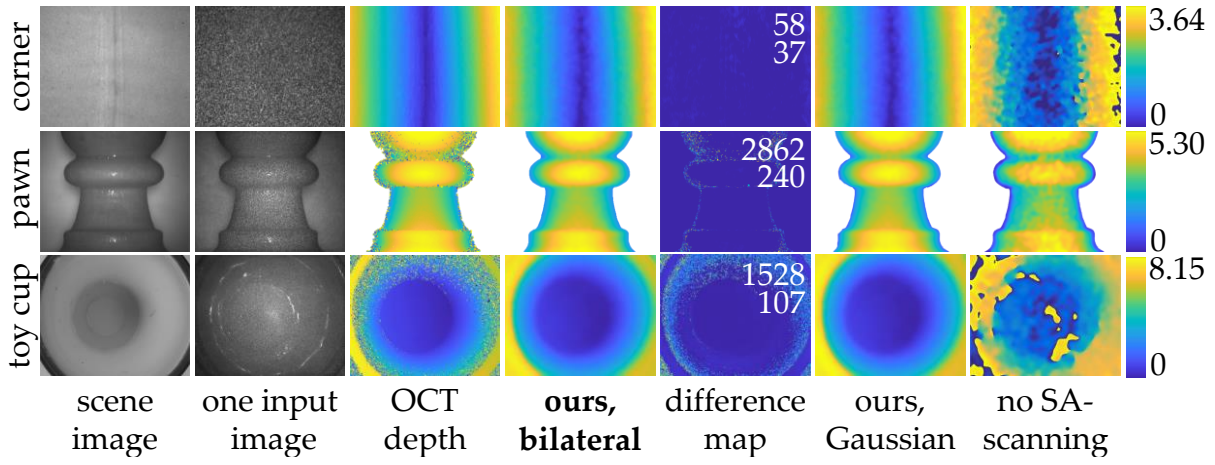


Figure 4.9: **Comparing swept-angle SWI and full-field OCT in macroscopic scenes** (from Figure 4.6). The difference maps show the absolute difference between recovered OCT and swept-angle SWI depths, and report the root-mean-square (top) and mean absolute (bottom) differences between the two. The OCT depths are captured at a resolution comparable to swept-angle SWI, and should not be considered ground truth. The colorbars are in millimeters.

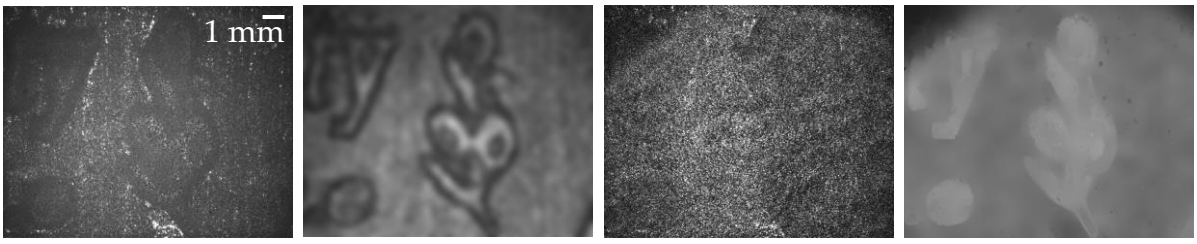


Figure 4.10: **Effect of swept-angle scanning**. Swept-angle illumination helps mitigate speckle noise and subsurface scattering effects, as exemplified in these example envelope images from the soap scene.

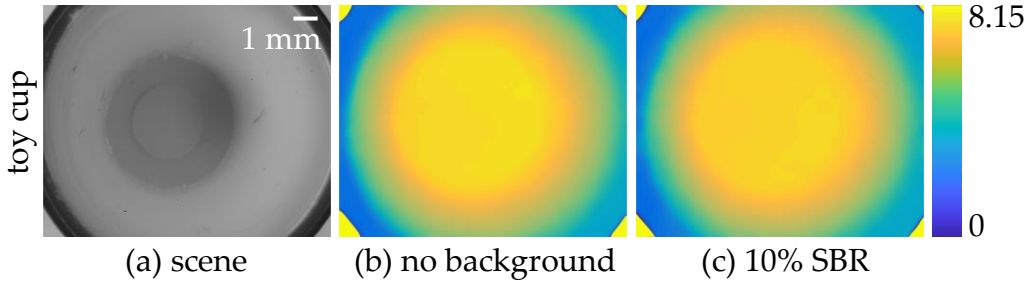
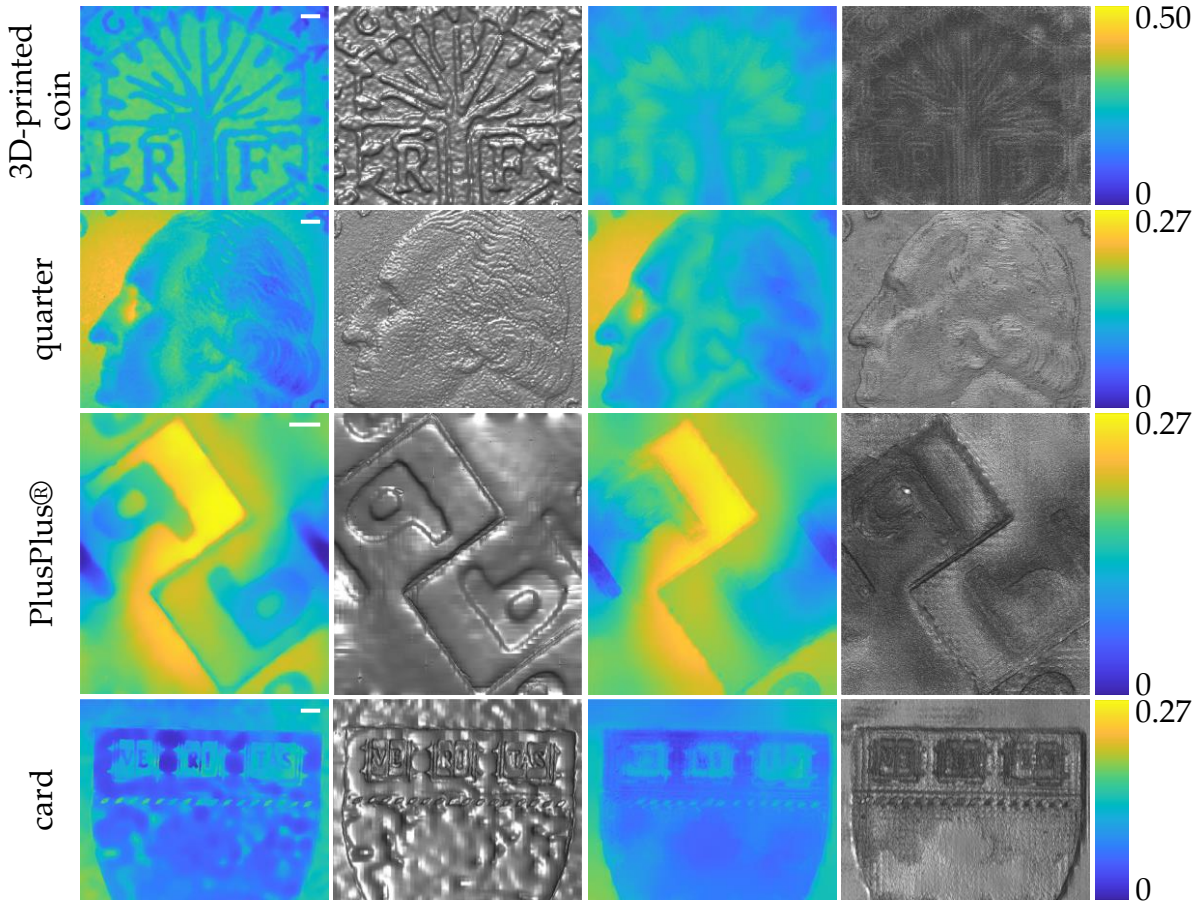


Figure 4.11: **Robustness to ambient light**. In (c), we shine external light on the sample so that the signal-to-background ratio (SBR) our laser illumination to ambient noise is 0.1. This greatly decreases the contrast of the interference speckle pattern. Despite this, there is little degradation in the quality of our recovered depth.

quality of full-field OCT, while being 50× faster. Figures 4.7 and 4.9 compare the performance of swept-angle SWI with full-field OCT implemented as in Gkioulekas et al. [18]. We also depict differences between swept-angle SWI and OCT depth reconstructions.

Indirect light rejection To demonstrate the difference swept-angle scanning makes, in Figure 4.10 we visualize one of the input data frames and one of the computed envelope images. We note that the spatial incoherence introduced by the swept-angle scanning significantly reduces interference speckle in the input camera images. This reduction is due to swept-angle scanning rejecting most indirect illumination arising from subsurface scattering. In addition, the envelope image computed with swept-angle scanning maintains the correct troughs of the engraving, while turning off the swept-angle scanning spatially blurs the envelope image across the troughs.



(a) full-field SWI (ours) (b) scanning+upsampling

Figure 4.12: **Comparison with upsampled scanning SWI**. We emulate scanning SWI by downsampling our swept-angle SWI depth. We then bilaterally upsample it to the original resolution.

Comparison with upsampled scanning SWI. A direct comparison of swept-angle SWI with scanning SWI is challenging, because of the differences between scanning and full-field setups (Figure 4.3, (b) versus (a), (c)). To qualitatively assess their relative performance, in Figure 4.12 we use the following experimental protocol: We downsample the depth map from our technique, by a factor of 35 in each dimension. This is approximately equal to the number of depths points a scanning SWI system would acquire at the same total exposure time as ours, when equipped with beam-steering optics that operate at a 30 kHz scan rate. We then use joint bilateral upsampling [32] with the same guide image as in our technique, to upsample the downsampled depth map back to its original resolution. We observe that, due to the sparse set of points scanning SWI can acquire within the exposure time, the reconstructions miss fine features such as the hair on the quarter and letters on the business card.

Robustness to ambient light. In Figure 4.11, we demonstrate the robustness of our method to ambient light on the toy cup scene. We shine a spotlight on the scene such that the signal-to-background ratio (SBR) of the laser illumination to ambient light is 0.1. Ambient light adds to the intensity measurement at the camera, but not to interference, thus reducing interference contrast and potentially degrading the depth reconstruction. However, we see that at this SBR, the depth recovered from the toy cup scene (Figure 4.11 (d)) is virtually unchanged from the depth recovered without ambient light (Figure 4.11 (c)). In addition, to reject ambient light, we can use an ultra-narrow spectral filter centered at the illumination wavelength, an advantage OCT lacks because of its broadband spectrum.

Table 4.2: **Quantitative evaluation of the resolution of our method**. MAE is the mean absolute error between measured and ground truth depths, MedAE is the median absolute error and RMSE is the root mean square error. All quantities are stated in µm.

type	kernel width	with swept-angle		w/o swept-angle	
		RMSE	MedAE	RMSE	MedAE
micro-scopic (period 400 µm)	7	8.2	4.8	18.9	13.2
	15	5.1	3.6	11.2	9.5
	21	2.0	1.6	10.5	7.3
	30	1.6	1.0	11.1	6.7
macro-scopic (period 16 mm)	7	471.4	300.3	1267.2	1351.0
	15	167.1	120.5	577.9	501.9
	21	78.7	50.9	609.5	412.2
	30	81.7	49.6	605.7	334.4

Trade-off between acquisition time and depth quality. The theoretical minimum number of measurements needed to reconstruct depth using SWI is 9 ($M = N = 3$). However, increasing M and N makes the depth reconstruction robust to measurement and speckle noise, yielding higher quality depth. There is, therefore, a trade-off between number of measurements $M \cdot N$ and depth quality.

Besides number of measurements, another factor contributing to acquisition time is the MEMS mirror scan we perform to create swept-angle illumination. If we decrease acquisition time, giving the mirror lesser time to complete one full scan of the source in the focal plane of the collimating lens, the spatial density of scanned points on the simulated emission area decreases. This reduces the effectivity of rejecting indirect light, and therefore reduces depth quality. This, again, creates a trade-off between acquisition time and depth quality.

In Figure 4.13, we demonstrate the effect of both these factors on depth quality. On the horizontal dimension, we use different $\{M, N\}$-shift phase retrieval algorithms, and on the vertical dimension, we use different per-image acquisition times as dictated by the scan time of the MEMS mirror. We observe that using higher M and N allows us to reduce the per-image acquisition time by requiring a lower scanned source density for equal visual depth quality. In particular, the 100 ms scan with the $\{4, 5\}$-shift algorithm performs as well as the 10 ms scan with the $\{5, 5\}$-shift algorithm, allowing us to reduce total acquisition time from 2 s to 250 ms.

Axial resolution. To quantify the axial resolution of our technique, we perform the following experiment. We use a second nanometer-accurate translation stage to place the chocolate scene from Figure 4.6 at different depths from the camera, at increments of 1 µm. We choose this scene because it is characterized by strong sub-surface scattering. We then compare how well full-field SWI performs in four categories: with microscopic ($\lambda_s = 400$ µm) or macroscopic ($\lambda_s = 16$ mm) synthetic wavelengths, and with or without swept-angle scanning. We perform this experiment using Gaussian filtering with kernels of different sizes, to additionally quantify lateral resolution. The results in Table 4.2 show that, in the microscopic case, we can achieve an axial resolution of approximately 5 µm and 1 µm, for kernel sizes 7 µm and 30 µm, respectively. In the macroscopic case, we can achieve an axial resolution of approximately 50 µm, for kernel sizes of 21 µm. We can trade off lateral for axial resolution, by increasing kernel size during phase retrieval. Figure 4.6 shows that this trade-off can become more favorable using bilateral filtering.

We additionally show the data we captured to assess our depth accuracy in Table 4.2.

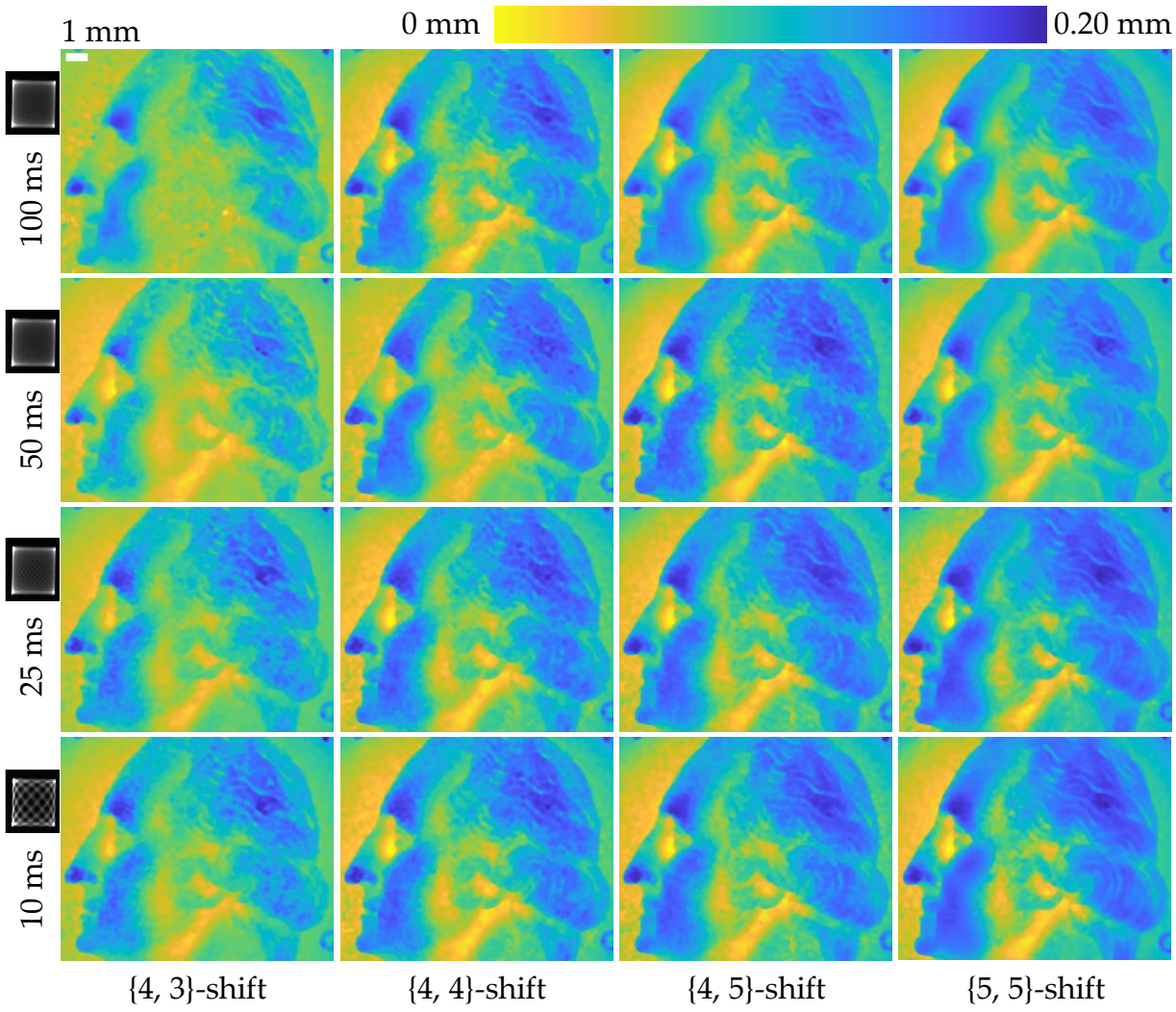


Figure 4.13: **Depth quality and acquisition time**. We show qualitatively the effects of per-image acquisition time (dictated by the period of the MEMS mirror scan) and the number of images acquired $M \cdot N$ on the quality of the reconstructed depth. The pattern in the black box next to per-image times is the scanned emission area in the focal plane of the illumination lens.

Figure 4.14 plots, on top, the estimated SWI depth relative to the ground truth depth provided by the scene translation stage. Figure 4.14(a) is captured with the swept-angle mechanism, whereas Figure 4.14(b) is captured with the mechanism off. Comparing the two figures, we see that the measured depth correlates with the groundtruth depth a lot stronger when we use swept-angle illumination versus when we do not.

Figure 4.14(c) and Figure 4.14(d) respectively show the same experiment performed at a *macroscopic* synthetic wavelength of 16 mm, the same as in Figure 4.8. These measurements also depict that swept-angle scanning is critical for micron-scale depth sensing.

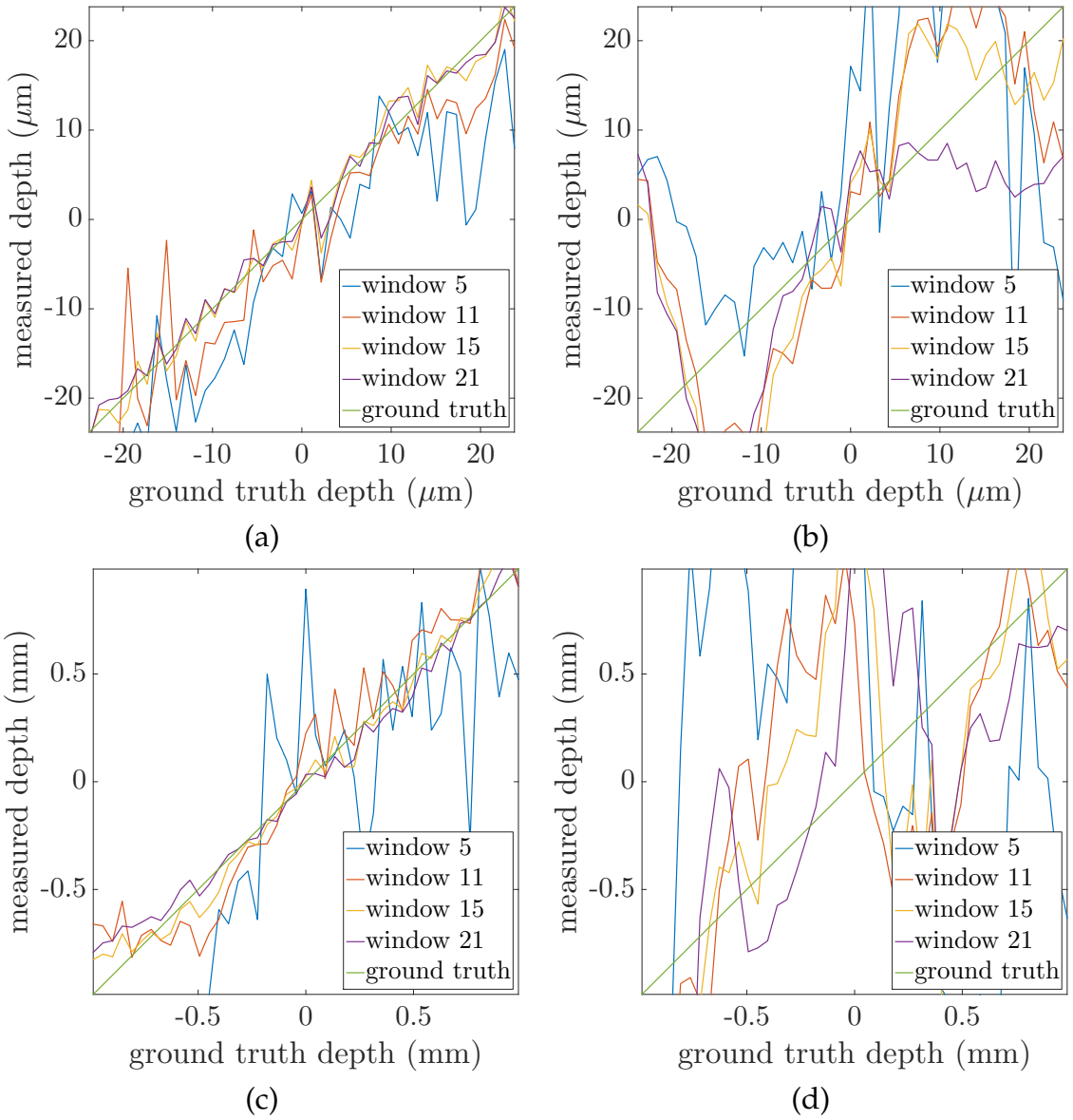


Figure 4.14: **Testing the depth resolution of our method**. We place the chocolate scene from Figure 4.6 at different distances from the camera using a translation stage and capture measurements using our method at each position. We do this under four conditions: **(a)** microscopic synthetic wavelength with swept-angle, **(b)** microscopic synthetic wavelength without swept-angle, **(c)** macroscopic synthetic wavelength with swept-angle, and **(d)** macroscopic synthetic wavelength without swept-angle. In each case, we plot the depth measured by our method against the ground truth position of the scene provided by the translation stage. The window parameter in the plots is the size of the Gaussian blur kernel.

4.5 Discussion

Phase wrapping. As discussed in Section 4.1, because SWI computes depth from phase, it only determines depth up to an integer multiple of the synthetic wavelength λ_s. Equivalently, all depths in the scene are *wrapped* to $[0, \lambda_s/2]$, even if the depth range is greater. To mitigate phase wrapping and extend the unambiguous depth range, it is common to capture measurements at multiple synthetic wavelengths, and use them to *unwrap* the phase estimate [3, 8, 9, 13]. Whereas, theoretically, two synthetic wavelengths are enough to uniquely determine all depths, in practice phase unwrapping techniques are very sensitive to measurement noise [21]. Combining phase unwrapping algorithms with our technique is an interesting future direction.

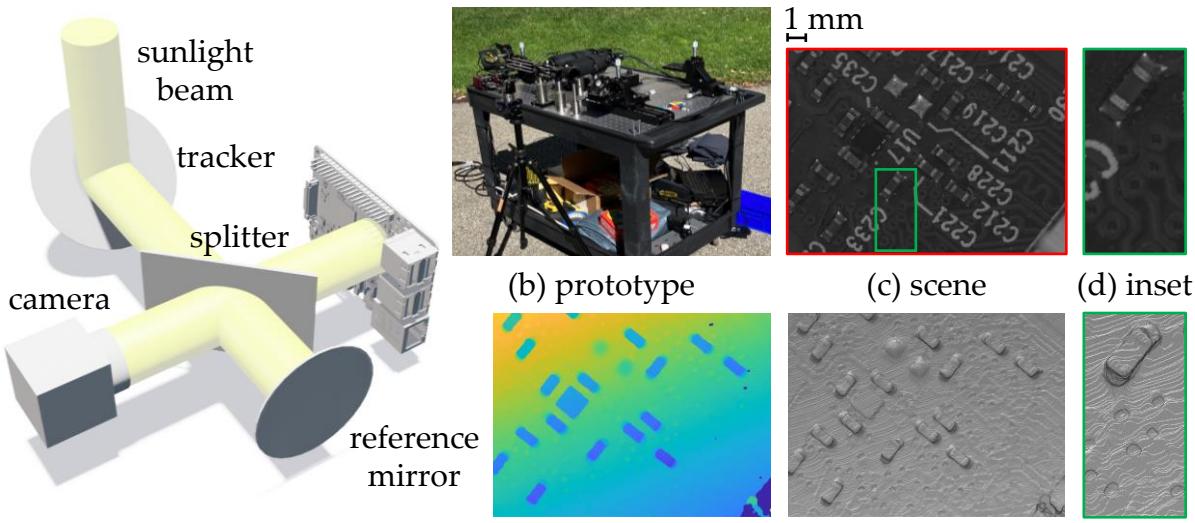
Toward real-time operation. The prototype in Figure 4.3(a) acquires measurements at a frame rate of 5 Hz, due to the need for 10 ms per-frame exposure time. If we use a stronger laser to reduce this time, the main bottleneck towards faster operation will be the need to perform phase shifts by physically translating the reference mirror. We can mitigate this bottleneck using a faster translation stage (for example, fast microscopy stages). Alternatively, we can eliminate this bottleneck by using phase modulators and multiple mirrors, to perform the shifts without mechanical motion. We leave the development of such a prototype as future work.

4.6 Summary

We presented a technique for fast depth sensing at micron-scale lateral and axial resolutions. Our technique, swept-angle synthetic wavelength interferometry, combines the complementary advantages of full-field interferometry (speed, pixel-level lateral resolution) and scanning interferometry (robustness to aberrations and indirect illumination). We demonstrated its advantages by scanning a variety of scenes with fine geometric features and strong subsurface scattering. We expect that our results will facilitate applications of swept-angle SWI in areas such as biomedical imaging, and industrial fabrication and inspection (Figure 4.2).

Chapter 5

Passive interferometry



(a) passive time-of-flight setup (b) prototype (c) scene (d) inset (e) depth map (f) rendered depth (g) inset

Figure 5.1: **Using sunlight interferometry to passively reconstruct part of a circuit board**. **(a)** and **(b)** show a schematic and photograph of the sunlight interferometer we build for passive time-of-flight imaging. We use this system to reconstruct part of a Raspberry Pi circuit board that has multiple resistors, soldering pads, and tracks. **(c)** and **(d)** show a picture of the scene as seen through the imaging camera, along with an inset highlighting fine geometric features. **(e)** shows the estimated depth map, and **(f)** and **(g)** the corresponding rendered 3D surface. We note that our technique manages to reconstruct very fine features such as the PCB tracks and through-holes, despite operating outdoors under adverse environment conditions.

So far in this thesis, we have considered interferometry setups that use active illumination we create using light emitted by broadband or single-frequency lasers. In

addition, since interferometry is notoriously sensitive to vibrations, the setups themselves were mounted upon vibration-isolated optical tables in a dark room for optimal interference SNR. These requirements make interferometry impractical for use outdoors, in the presence of strong ambient illumination that overwhelms the light source, or in power-constrained applications. By contrast, passive 3D sensing techniques, such as multi-view stereo, shape from shading, and depth from (de)focus, use ambient light from the environment itself as illumination. However, all of these techniques can achieve resolutions in the order of hundreds of micrometers at best, placing them out of scope for applications requiring micron-scale resolutions.

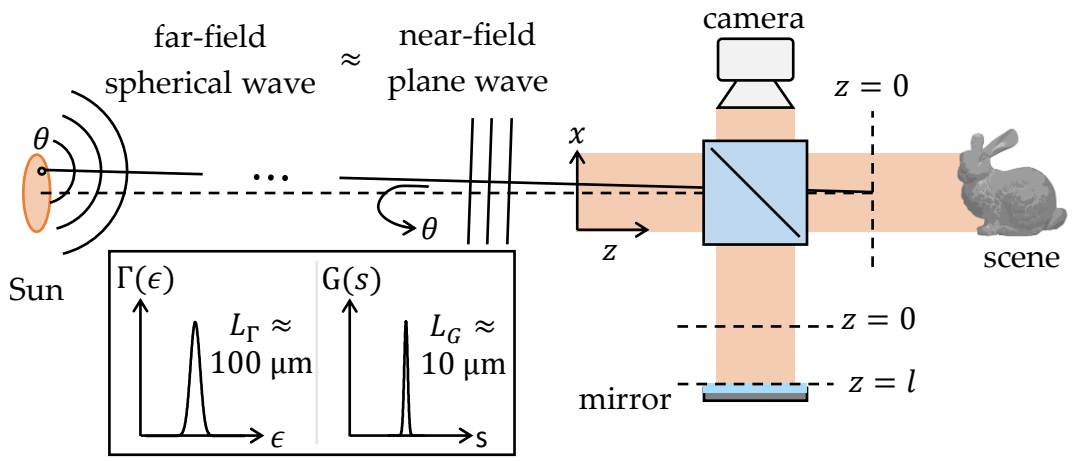
In this chapter, we introduce a completely passive interferometric imaging technique for micron-scale 3D sensing. We focus upon imaging using the most ubiquitous passive light source available: sunlight. We explain how full-field interferometry with sunlight illumination provides us with impulse time-of-flight measurements robustly to severe indirect illumination effects such as interreflections and subsurface scattering, as shown in Figure 4.4.

To demonstrate the capabilities of passive interferometry, we prototype a full-field Michelson interferometer, modified to use sunlight as the only light source. We operate our prototype outdoors, under direct sunlight and adverse experimental conditions (wind, machine vibrations, vehicle traffic). We show depth reconstruction results with micron axial resolutions and pixel-level lateral resolutions, that are competitive with the ones acquired with active illumination inside the lab on a vibration-isolated optical table. Figure 5.1 illustrates the exquisite depth quality passive sunlight interferometry provides for one of our reconstruction results in the face of the challenges mentioned above.

We further use sunlight interferometry for computational light transport tasks such as direct-only imaging and depth sensing through optically-thin scattering layers. More broadly, our findings open the door for the deployment of interferometric techniques in uncontrolled outdoor environments, and for the development of passive computational light transport capabilities such as the ones we demonstrated in Chapters 3 and 4.

5.1 Coherence properties of sunlight

As in previous chapters, we theoretically analyze the coherence properties of sunlight to assess the probing capabilities of sunlight interferometry. Since the Sun is an extended, spatially incoherent source of light, each point on the sun serves as one of the incoherent modes of illumination in Equation (2.1). This allows us to perform spatial probing. In



(a) sunlight illumination (b) full-field Michelson interferometer

Figure 5.2: **Coordinate system used for the analysis in this chapter. (a)** Since the sun is an area source, it subtends a finite angle at the input plane $z = 0$. Each point θ on the sun emits a spherical wave towards the interferometer. **(b)** Since the sun is in the far-field of the interferometer input, the spherical wave looks approximately like a plane wave traveling at an angle θ to the optical axis of the interferometer.

addition, the Sun emits broadband light, allowing us to perform temporal probing.

We use the coordinate system depicted in Figure 5.2 for our analysis. We parameterize the modes of the illumination by labeling points on the Sun by the angles θ they subtend to the optical axis of the interferometer. The Sun spans an angle Θ in the sky, lying within the angles $\theta \in [-\Theta/2, \Theta/2]$ if centered along the optical axis of the interferometer. Since the Sun lies in the far-field of the interferometer input, the wave created by a point on the Sun is paraxially a plane wave at the interferometer. For a point subtending an angle θ to the optical axis, this is a plane wave traveling at angle θ to the optical axis. The wave field of this illumination at the input plane $z = 0$ can be written as

$$\boldsymbol{u}_i(x, z; \theta) = \exp\left(-\mathrm{j}\kappa\left(x \sin\theta + z\cos\theta\right)\right). \tag{5.1}$$

The Sun subtends $\Theta \approx 0.5^\circ$ of angular extent in the sky. Under this angular extent, the approximation $\sin\theta \approx \theta$ holds, allowing us to re-write the field as:

$$\boldsymbol{u}_i(x, z; \theta) = \exp\left(-\mathrm{j}\kappa\left(x\theta + z\right)\right). \tag{5.2}$$

This is very similar to the case of transmission probing with coded mutual intensity with an emitter as in Equation (3.21): Indeed, Fraunhofer diffraction tells us that placing

an emitter in the focal plane of a collimating lens has the same effect as placing an appropriately scaled version of the emitter in the far-field. Accounting for the finite size of the Sun in the sky, we can write the field as

$$\boldsymbol{u}_i\left(x, z; \theta\right) = \Pi_\Theta\left(\theta\right) \exp\left(-\mathrm{j}\kappa\left(x\theta + z\right)\right), \tag{5.3}$$

where $\Pi_\Theta\left(\theta\right)$ is the rectangular function, which is 1 for $\theta \in [-\Theta/2, \Theta/2]$ and 0 otherwise. Using this plane wave profile and the spectrum $S\left(\kappa\right)$ of sunlight, we can write the mutual intensity function from Equation (2.16) as:

$$\begin{aligned}\Gamma\left(x, x'\right) &= \int_\theta \Pi_\Theta^2\left(\theta\right) \exp\left(-\mathrm{j}\kappa\left(x'\theta + z\right)\right) \left[\exp\left(-\mathrm{j}\kappa\left(x\theta + z\right)\right)\right]^* \, \mathrm{d}\theta \\ &= \int_\theta \Pi_\Theta\left(\theta\right) \exp\left(-\mathrm{j}\kappa\left((x' - x)\theta\right)\right) \, \mathrm{d}\theta,\end{aligned} \tag{5.4}$$

where we have used the property that $\Pi_\Theta^2\left(\theta\right) = \Pi_\Theta\left(\theta\right)$. This depends only on $\epsilon = x' - x$, and therefore we can re-write it as

$$\begin{aligned}\Gamma\left(\epsilon\right) &= \int_\theta \Pi_\Theta\left(\theta\right) \exp\left(-\mathrm{j}\kappa\epsilon\theta\right) \, \mathrm{d}\theta \\ &= \Theta \operatorname{sinc}\left(\Theta\kappa\epsilon\right)\end{aligned} \tag{5.5}$$

The spatial coherence length L_Γ of the illumination is the width of this sinc:

$$L_\Gamma \approx {}^{1.2}/_{\Theta\kappa}. \tag{5.6}$$

As usual, the temporal coherence function of sunlight is the Fourier transform of the spectrum of the incoming illumination. Its width is the temporal coherence length $L_\mathcal{G}$. Previous work has estimated the temporal coherence length of direct sunlight at sea level to be approximately $L_\mathcal{G} = 0.3\,\mathrm{nm}$ [59]. However, performing sunlight interferometry at these axial resolutions using the full spectrum of sunlight is highly impractical: First, chromatic aberrations from optics (lenses, beamsplitters) when performing interferometry with very broadband illumination can drastically reduce interference contrast [15]. Second, environment vibrations when operating an interferometer outdoors are in the order of hundreds of micrometers. Together, these two factors result in impractically low SNR. Instead, we limit spectral bandwidth using a spectral filter with a Gaussian transmission profile centered around 550 nm with bandwidth 20 nm. The temporal coherence length $L_\mathcal{G}$ is now the width of the Fourier transform of this transmission profile.

With our acquisition procedure, which we will explain next, we measure the squared magnitude of the correlation and therefore perform light transport probing. As we saw in Section 3.2, for sufficiently diffuse scenes, measuring the squared correlation magnitude probes the transmission function by the squared magnitude of the coherent probing function $|\mathcal{P}^c(\epsilon, s)|^2$ (Equation (3.16)):

$$|\mathcal{C}(x; \mathcal{P}^c)|^2 \approx \frac{1}{c} \int_{\epsilon,s} \mathcal{T}(x, x+\epsilon, {}^{s+2l}/_c) \, |\Theta \operatorname{sinc}(\Theta\kappa\epsilon)|^2 \, |\mathcal{G}(s)|^2 \, \mathrm{d}\epsilon \, \mathrm{d}s. \tag{5.7}$$

For an infinitely large and broadband Sun, this expression measures

$$|\mathcal{C}(x; \mathcal{P}^c)|^2 = \mathcal{T}(x, x, {}^{2l}/_c), \tag{5.8}$$

performing exactly direct-only transient imaging [18, 51] by scanning the reference arm position l. A finite-sized and spectrally bandlimited Sun blurs the probed transport function across the spatial and temporal dimensions with a blur kernel of widths given by L_Γ and $L_\mathcal{G}$ respectively. Then, the lateral resolution of sunlight interferometry is limited by the spatial coherence length, and the axial resolution is limited by the temporal coherence length. As we will see in Section 5.4 and Figure 5.7, the properties of sunlight set both of these to allow us to perform micron-scale time-of-flight imaging.

5.2 Direct-only time-of-flight imaging

We now use the results of Section 5.1 to derive a procedure for performing direct-only time-of-flight imaging using sunlight interferometry. We assume that we want to scan a depth range equal to D. We discretize this depth range to a set of axial locations $l_{m+1} = l_m + {}^{L_\mathcal{G}}/_2$, $m = 1, \ldots, M$, where $M = \lceil {}^{2D}/_{L_\mathcal{G}} \rceil$. The axial separation between consecutive axial locations is due to the fact that, following the discussion at the end of the previous section, the temporal coherence length $L_\mathcal{G}$ determines the lateral resolution.

We use the translation stage of our setup to move the reference mirror to each of the axial locations l_m. At each location, we use the camera to capture an image $I_m(x)$ – as we use a two-dimensional sensor, we acquire each such image for all locations x in the field of view in a single exposure. We assume that, from each image $I_m(x)$, we can isolate the corresponding correlation $|\mathcal{C}_m(x)|^2$ term in Equation (5.7), as we explain next.

Computing correlation. The correlation $\mathcal{C}_m(x)$ is complex: this suggests that we can use phase shifting [12] to estimate both its amplitude and phase. However, this would

require shifting the reference mirror by fractions of the optical wavelength $\bar{\lambda} = 1/\bar{\kappa}$ and capturing 3-4 images around *each* location l_m. This makes acquisition prohibitively long, and is also impractical when operating outdoors – environment vibrations make it difficult to accurately perform sub-wavelength shifts.

Instead, we use the procedure proposed by Gkioulekas et al. [18] to approximate the squared correlation amplitude $\mathcal{T}\left(x, x, 2l/c\right)$: First, we estimate the interference-free image $\mathbf{i}_m\left(x\right)$ by averaging nearby frames in the axial scan,

$$\mathbf{i}_m\left(x\right) \approx 1/N \sum_{n=m-N/2}^{N/2} I_n\left(x\right). \tag{5.9}$$

Then, we estimate the squared interference as

$$\mathrm{Re}\left\{\mathcal{C}_m\left(x\right)\right\}^2 \approx \left(I_m\left(x\right) - \mathbf{i}_m\left(x\right)\right)^2 \equiv \mathbf{r}_m\left(x\right). \tag{5.10}$$

Lastly, we estimate the squared correlation amplitude as

$$\tilde{\mathcal{T}}\left(x, x, 2l_m/c\right) = \left(G_p * \mathbf{r}_m\right)\left(x\right), \tag{5.11}$$

where G_p is a Gaussian kernel with standard deviation of p pixels. We visualize our pipeline for correlation estimation in Figure 5.3. The intuition behind the blurring operation in Equation (5.11) is as follows: When imaging scenes with rough surfaces, the interference appears as speckle. This speckle is created due to pseudo-random sub-wavelength pathlength shifts, accounting for the surface's microstructure. Blurring (squared) interference speckle around some location is approximately equal to averaging squared interference measurements at sub-wavelength shifts. This is how phase shifting techniques estimate squared correlation amplitude [12].

Finally, we then estimate depth for each point in the scene as:

$$d\left(x\right) = \mathrm{argmax}_{l_m}\, \tilde{\mathcal{T}}\left(x, x, 2l_m/c\right). \tag{5.12}$$

We note that Equation (5.12) assumes that the direct-only transient is has a single peak, corresponding to a single reflecting surface at each location x. This assumption is violated, for instance, when the camera observes a semi-transparent object in front of another object. However, in practice, our procedure can handle such cases without issue: As we record the entire transient function $\mathcal{T}\left(x, x, 2l_m/c\right)$, we can search for multiple peaks and detect the depth of all reflecting surfaces. We show this experimentally in Section 5.4.

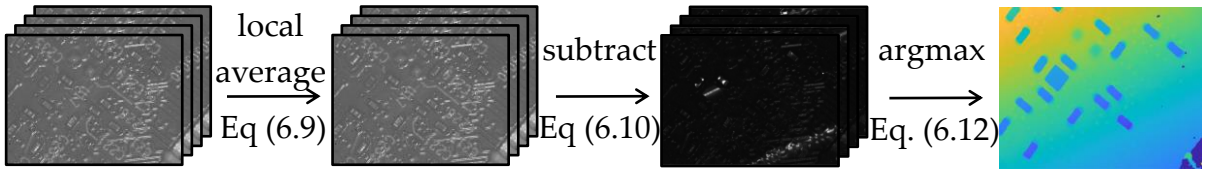


Figure 5.3: **Processing pipeline.** **(a)** We capture a stack of images with the reference arm placed at a dense set of positions. **(b)** We blur the stack temporally and spatially, to estimate interference and correlation amplitude. **(c)** We extract a depth map by detecting the axial location maximizing correlation amplitude.

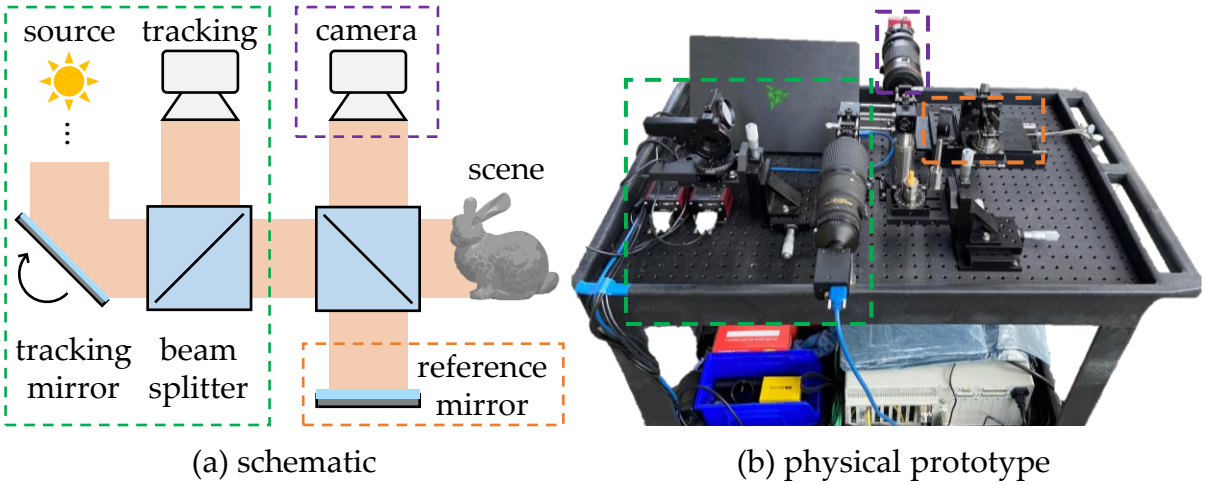


Figure 5.4: **(a) Schematic and (b) physical prototype** of the sunlight interferometry setup on a utility cart. In order to be able to do interferometry outdoors, we build the optical setup on a passively-damped breadboard placed on an ordinary utility cart. We use a tracking assembly that uses a mirror whose position is controlled with a closed-loop system to inject ambient light into the interferometer, as shown in the green box.

5.3 Hardware prototype

We build an experimental prototype to demonstrate the capabilities of sunlight interferometry. We show an image of our prototype in Figure 5.4. We discuss here some key implementation aspects, and refer to the supplement for full implementation details, including a parts list, construction, alignment, and calibration procedures.

Key specifications. Our prototype uses a camera equipped with a CCD sensor (pixel pitch 3.7 µm, 3400 × 2700 pixels) and a compound lens (focal length 300 mm). We use a reproduction ratio of 1:1, resulting in a field of view of 12.5 mm × 10 mm, and working distance of 600 mm. For axial scanning, our prototype uses a translation stage with

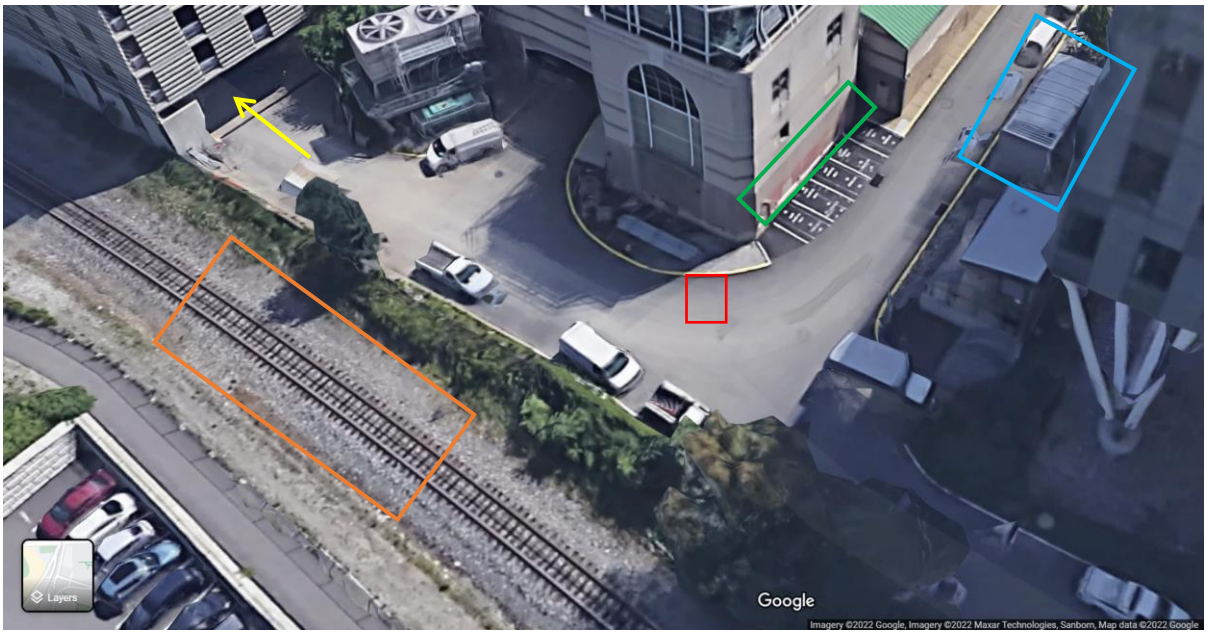
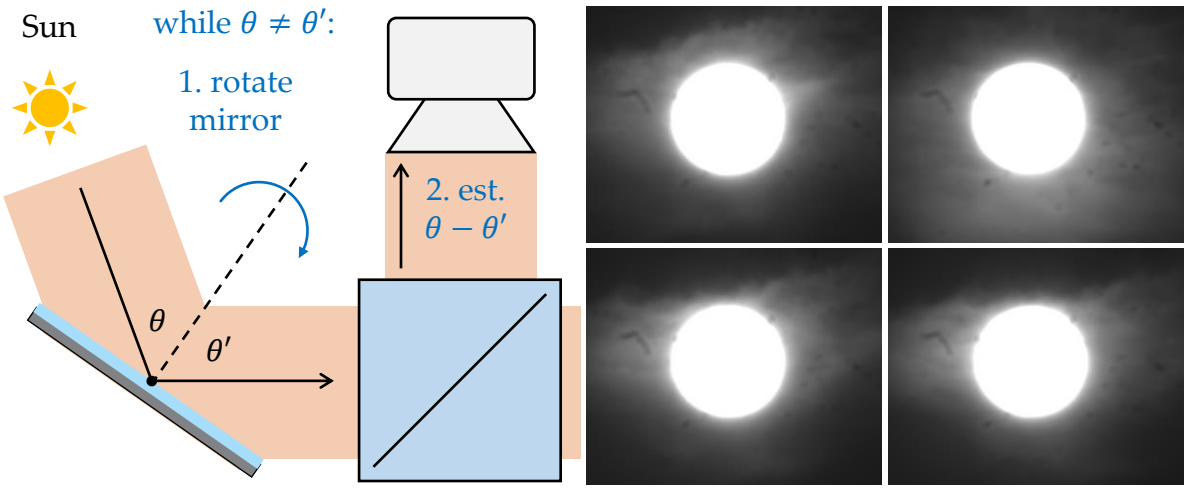


Figure 5.5: **Google Maps view of experimental site.** This image showcases the challenging conditions of our experiments. **Red**: setup location. **Yellow**: parking garage, causing constant traffic with multiple vehicles (including trucks) passing per acquisition. **Orange**: train tracks, with trains at frequent intervals. **Green**: building air intake causing a constant loud background hum. **Blue**: power transformer also causing a loud constant background hum.

resolution 10 nm. We use an axial stepsize of 5 μm (Section 5.1) and capture 1000 images per scan at exposure times of 50 ms. This results in acquisition times around a minute per scene. We follow Gkioulekas et al. [18] in setting the various imaging settings to maximize interference contrast, and thus signal-to-noise ratio (SNR).

Tracking the Sun. As it is not practical to orient the optical axis of the interferometer towards the Sun, and to account for motion of the Sun during acquisition, we equip our setup with a custom Sun-tracking assembly (Figure 5.4, green boxes). This assembly comprises a mirror whose 3D orientation is controlled via two motorized rotation stages. We set the mirror orientation using information from a tracking camera separate from the imaging camera: We set the camera so that its optical axis is parallel to that of the interferometer, and its lens focused at infinity. The camera directly images the Sun through the mirror, and provides feedback (through a digital PID controller, Figure 5.6(a)) to the rotation stages to adjust the mirror orientation, so that the Sun remains at the center of the camera's field of view.



(a) tracking mechanism (b) tracking camera images

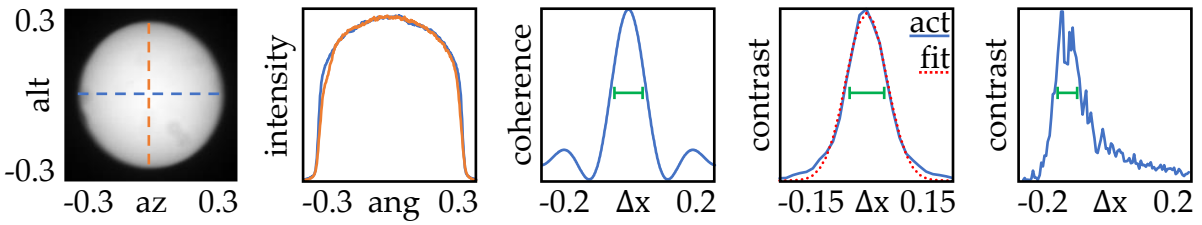
Figure 5.6: **Sun tracking. (a)** In a closed loop, we estimate the direction of the light beam reflecting from the mirror with a tracking camera focused at infinity, and correct it by rotating the mirror. **(b)** Tracking camera images of the Sun during a centering operation.

Outdoor operation. To perform experiments outdoors and mitigate the vibration effects, we build our setup on an optical breadboard mounted on a utility cart. The imaged scenes are on the ground or a tripod, and are fully immersed in ambient light. Figure 5.5 shows our experimental site. It includes strong sources of environment noise, including moving vehicles, air, and ambient sound. These experimental conditions are in stark contrast with those in prior works on interferometric computational imaging [10, 18, 33, 38, 39, 42], which require carefully-controlled environments (for example, dark room, vibration-isolated optical tables, no air flow).

Cloudy sky conditions. We performed most of our experiments under direct sunlight. However, we found that under scattered or very high-altitude clouds, we are still able to obtain high-quality depth, even if the clouds are fast-moving. Figure 5.6(b) shows frames acquired by the tracking camera during a single acquisition, showing the tracked Sun under fast-moving, scattered cloudy conditions.

5.4 Experiments

Measuring the temporal coherence length. To measure the resulting temporal coherence length, we perform two experiments where we place either a mirror or a planar diffuser at the scene arm of our setup. For each scene, we estimate $\mathcal{T}$ at a dense set of



(a) Sun tracking image (b) cross-sections (c) spatial coherence (d) temporal coherence (e) temporal coherence

Figure 5.7: **Measuring the coherence properties of sunlight**. **(a-b)** An image of the Sun taken by the tracking camera yields an angular size of 0.57°. **(c)** From this, we estimate the spatial coherence length of sunlight as $L_\Gamma \approx 100\,\mu\text{m}$. **(d-e)** With a 550±20 nm spectral filter, we measure the temporal coherence function of sunlight using a mirror and a diffuser respectively. From the full-width at half-maximum of a Gaussian fit to the mirror coherence function, we estimate the temporal coherence length to be approximately $L_\mathcal{G} = 10\,\mu\text{m}$. The x-axes in (c-e) are in mm.

reference axial locations. From Equation (5.7), we expect the square root of the measurements to be shaped like a Gaussian with standard deviation equal to the temporal coherence length. We perform a Gaussian fit to the square root of the measurements, and estimate the temporal coherence length as the fit's full-width half-maximum Figure 5.7 shows the results. For both mirror and planar diffuser, we estimate the temporal coherence length to be approximately $L_\mathcal{G} = 10\,\mu\text{m}$.

Measuring the spatial coherence length. To measure the spatial coherence length of sunlight, we use the tracking camera of our setup to capture a direct image of the Sun. At infinity focus, the spatial extent of the Sun in this image is directly proportional to the angular extent of sunlight (up to the focal length of the lens). We can then use Equation (5.5) to estimate the spatial coherence length. Figure 5.7 shows the results. We estimate the angular extent of sunlight to be around 0.57°, corresponding to a spatial coherence length of approximately $L_\Gamma = 100\,\mu\text{m}$. This is consistent with theoretical and experimental estimates in prior work [2, 43].

Depth scanning. In Figure 5.8, we show depth scans for variety of scenes. These feature micrometer-scale geometric details, and pose a variety of challenges including low reflectivity (circuit board), specularities (coin, pawn) and strong subsurface scattering (chocolate, soap, pill). We obtain high quality reconstructions, despite operating under very adverse environment conditions (Figure 5.4).

In addition, as Equation (5.8) shows, the function $\mathcal{T}\left(x, x, {}^{2l}/_{c}\right)$ we estimate from our

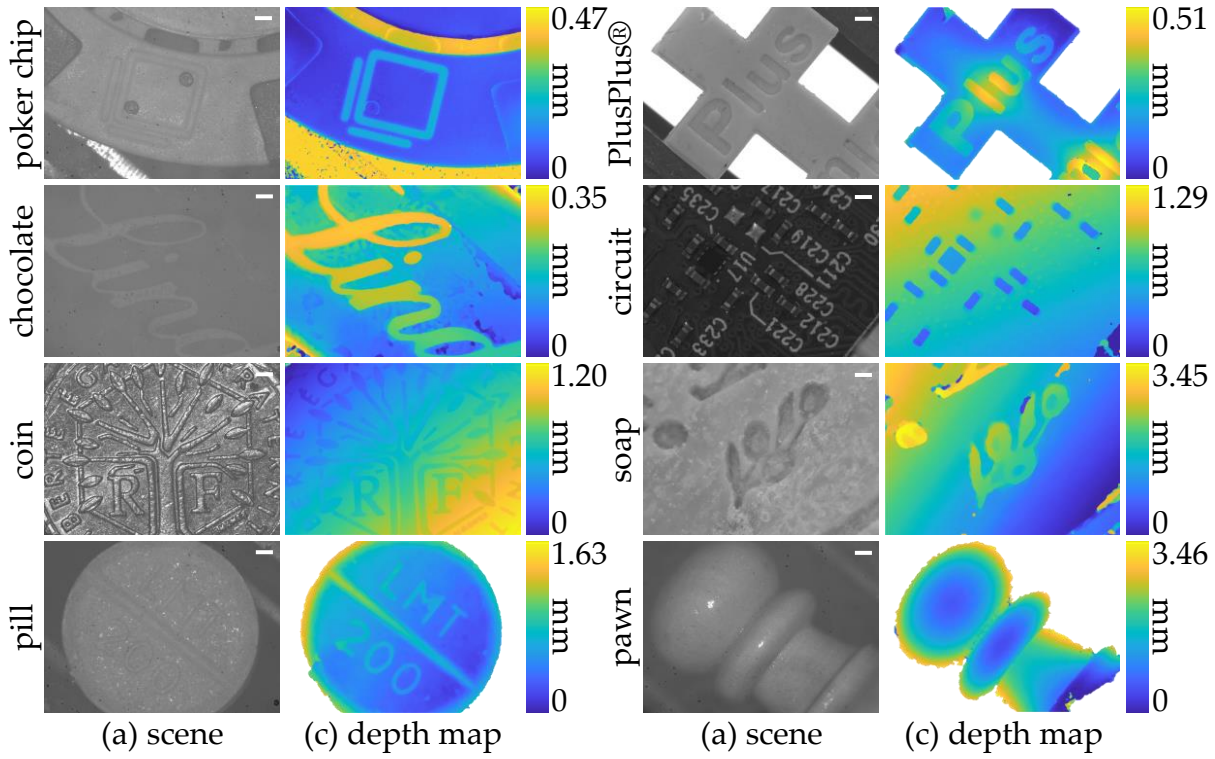


Figure 5.8: **Depth reconstructions for a variety of challenging scenes using sunlight**. The challenges include environmental factors (Figure 5.5) and low signal quality (detailed in the text). The white bars on the scene images indicate one millimeter of lateral size.

depth recovery pipeline is the direct-only transient response of the scene to incoming illumination. We can then visualize slices of the propagation of direct-only light in the scene as a stack of images of $\mathcal{T}$. We show some of these slices in Figure 5.9 for the scenes in Figure 5.8.

Scanning under extremely low SNR. A key challenge in sunlight interferometry is the very low SNR in captured images. This is for two reasons: First, in Equation (5.10), the interference $\mathbf{i}$ is typically much smaller than the interference-free image I. Second, the scenes are illuminated not only by the sunlight directed at them by the interferometer, but also by the ambient sunlight, further reducing interference contrast. To test the SNR limits of our technique, in Figure 5.10, we scan a semi-transparent object (gummy bear) that backreflects very little light. Despite this, our technique still acquires accurate depth for most of the scene, except for parts observed at near-grazing angles (no backreflected light). The transient frames in Figure 5.10(c) illustrate the extremely low SNR of the extracted interference.

Result with off-table scene. All the scenes in Figure 5.8 were captured with the scene placed on the optical table. In contrast, in Figure 5.11, we capture depth a scene that is placed on a tripod independently of the table, as shown in Figure 5.1(b). Imaging off-table scenes is more challenging because vibrations from the ground affect the scene and the interferometer independently. However, with some simple image processing, we are able to acquire depth comparable to when the scene is placed on the optical table. This result shows the potential of passive interferometry for use in robotics and manipulation applications.

Time-of-flight with indoor illumination. Lastly, even though our focus is on sunlight interferometry, Figure 5.12 shows an experiment testing the ability to perform time-of-flight imaging with passive *indoor* illumination. For this, we point the Sun tracking module of our setup to an ordinary desk lamp about five feet from the tracking mirror. We use as scene the same metallic coin as in Figure 5.8, which backreflects a lot of light, resulting in high SNR. We note, however, that this scanning setting remains challenging for non-metallic objects, because of the very low SNR between backreflected versus ambient light. Even though the lamp isn't in the far-field of the interferometer input, we obtain depth comparable with depth from sunlight interferometry. However, being in the near-field, every source point presents a quadratic phase profile across the illumination, distorting depth acquisition. We leave correcting this distortion to future work.

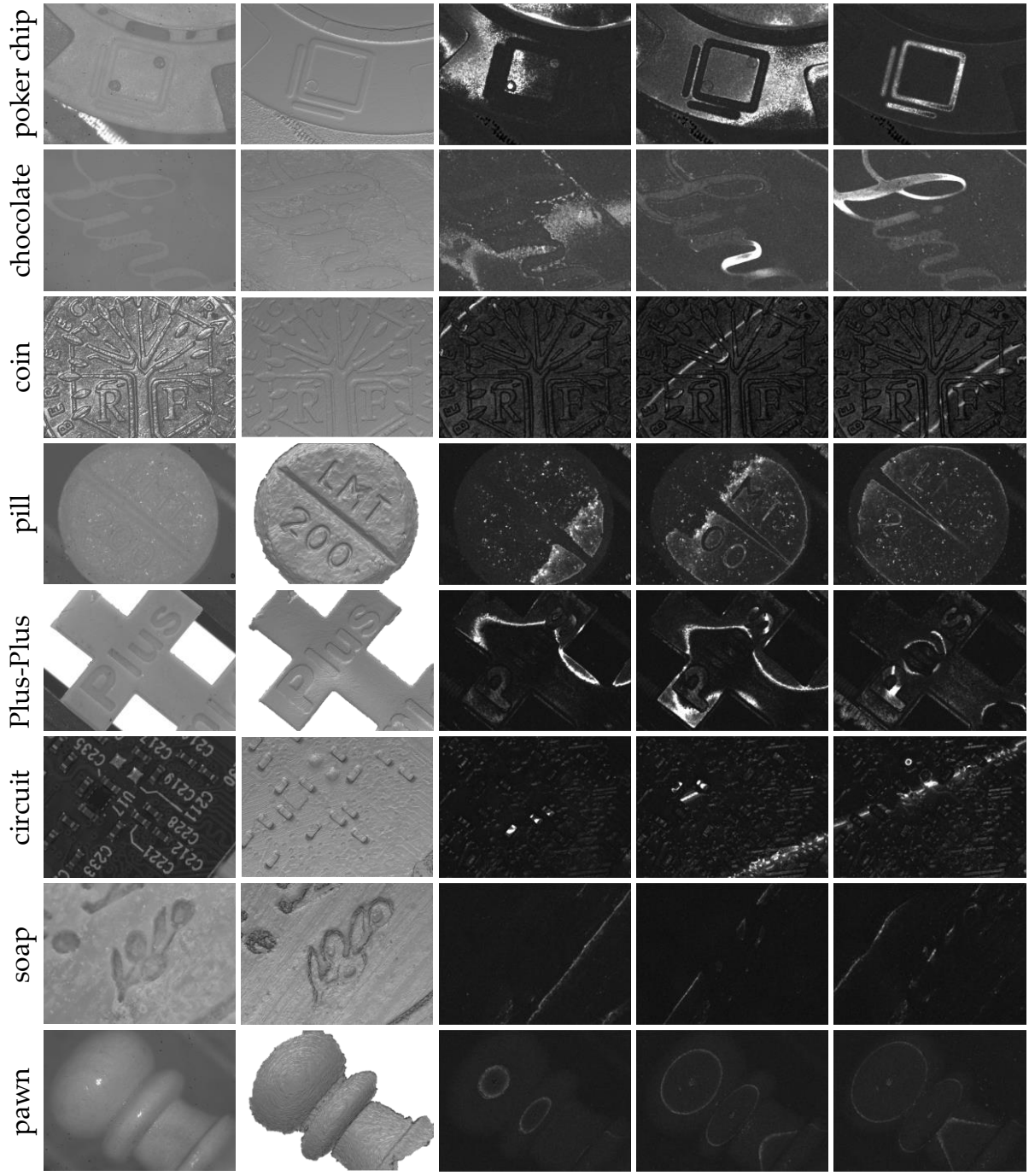
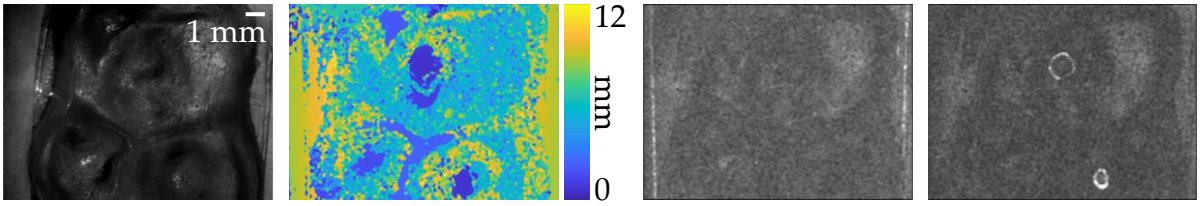
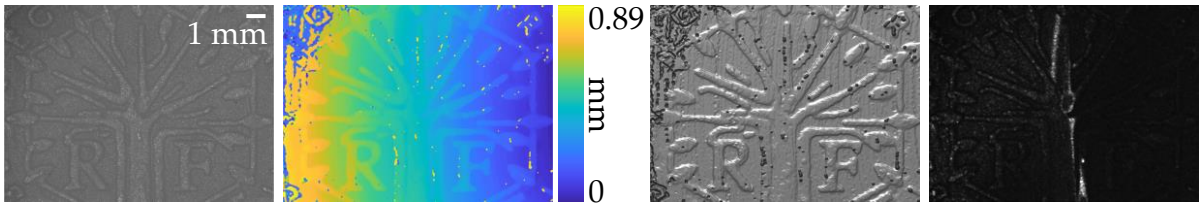


Figure 5.9: **Depth reconstructions for a variety of challenging scenes using sunlight**. The challenges include environmental factors (Figure 5.5) and low signal quality (detailed in the text). The white bars on the scene images indicate one millimeter of lateral size.



(a) scene (b) depth (c) transient video frames

Figure 5.10: **Depth reconstruction under extremely low SNR**. The gummy bear has very little backreflected direct light, making it extremely challenging to capture depth. The transient frames in (c) illustrate the extremely low SNR of the extracted interference.



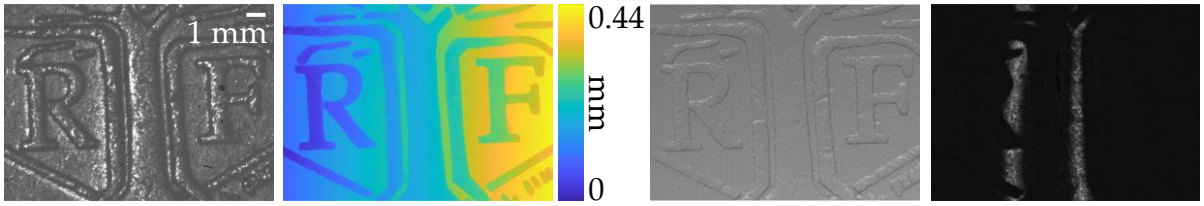
(a) scene (b) depth (c) rendered depth (d) one transient video frame

Figure 5.11: **Depth reconstruction for a scene placed off the optical table**. Though this is very challenging due to vibrations independent of the setup, we acquire depth with comparable quality with some simple signal processing.

Seeing through diffusers. To test the ability of our technique to isolate direct-only (ballistic) light and measure multiple depth peaks, in Figure 5.13, we scan a scene where a coin is occluded from the camera by a ground glass diffuser. The conventional images of the scene show strong blur because of the diffuser. Our technique can accurately acquire the depth *and* clean images of the occluded coin.

5.5 Summary

We have demonstrated a passive interferometric method for micron-scale depth sensing outside the controlled environment of an optics lab. Using light from ambient sources such as the Sun removed the need to use expensive, power-hungry, high-power or eye-unsafe lasers for interferometry. The inherent incoherence properties of sunlight made it possible to acquire time-of-flight information at micrometer scales, and insensitive to indirect illumination effects that confound depth inference. Additionally, engineering a robust optical setup mitigated most of the detrimental effects associated with performing experiments in uncontrolled outdoor environments, such as vibrations and wind, that



(a) scene (b) depth (c) rendered depth (d) one transient video frame

Figure 5.12: **Depth reconstruction with indoor illumination** created by an ordinary lamp at a distance of about five feet from the interferometer. Even though the lamp is not in the far-field of the interferometer, we can measure micron-scale depth.

typically hinder interferometric measurements. Together, these contributions allowed us to demonstrate, for the first time, passive imaging capabilities such as micrometer-scanning of challenging scenes, direct-only imaging, and imaging through diffusers. We also take first steps towards enabling these capabilities under passive *indoor* illumination. We hope that our work will motivate applications in areas such as inspection, fabrication, robotic grasping, and biomedical imaging.

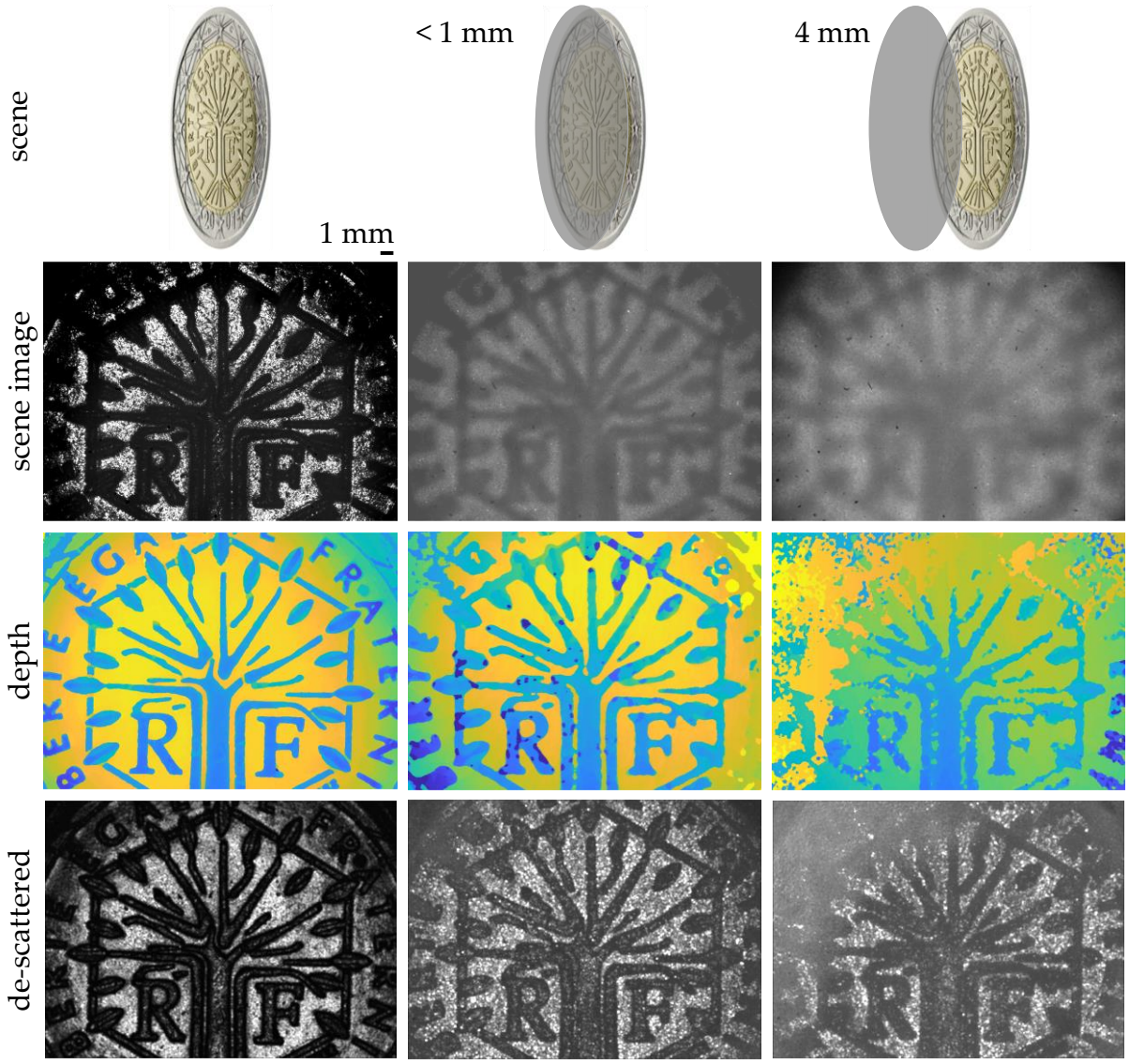


Figure 5.13: **Imaging and depth acquisition through diffusers**. We image a metallic coin under three conditions: **(a)** directly, **(b)** with a ground glass diffuser stuck to it and **(c)** with the diffuser placed 4 mm away. The images of the scene in **(b-c)** are blurred due to scattering in the diffuser. The interference component rejects scattering.

Chapter 6

Future work

HERE, we will discuss some future directions for both active and passive interferometric imaging. These directions would enable us to either relax the assumptions made in this thesis to achieve the full generality of the theory in Chapter 2; make implementations of projects in this thesis more robust and fast; or show further computational applications of measurements captured with interferometry. In summary, here are the projects we propose:

Complete spatio-temporal probing. In Chapter 3, we demonstrated just convolutional spatial transmission probing. In full generality, the theory in Chapter 2 allows us to probe the *time-resolved* transmission matrix. In addition, the assumption of convolutional probing can be overcome to implement *per-pixel* varying probing. Implementing this general form of probing can be split into four increasingly challenging projects:

1. *Temporal probing with spectrally coded interferometry.*
2. *Per-pixel probing.*
3. *Per-pixel spatio-temporal probing.*
4. *Passive probing.*

Computational imaging with interferometry. The measurements we acquire with our optical setups can be used further for computational imaging in diverse fields.

1. *Heterogeneous inverse scattering.* The capability of interferometry to probe light *transport* in highly-randomizing scenes makes it possible to perform inverse scattering as proposed by Gkioulekas et al. [19]. The applications of inverse scattering are numerous: characterizing materials, determining the three-dimensional optical

properties of heterogeneous scattering media and medical imaging, to name a few. The micron resolution capability of interferometry makes it uniquely suited for inverse scattering problems at micron scales.

2. *Neural processing of interferometry data.* Inspired by Che et al. [7], we can envision computational methods using our measurements being implemented with neural networks. Since our acquisition pipelines are differentiable, we can backpropagate through the network and our acquisition to jointly optimize acquisition and reconstruction.

6.1 Temporal probing with spectrally coded interferometry

We recall again Equation (3.3) for interferometric probing: We can probe the complex light transmission with the probing function $\mathcal{P}^c(\epsilon, s) = \Gamma(\epsilon) G(s)$ for a source spectrum $S(\kappa)$ and a mutual intensity profile $\Gamma(\epsilon)$. In analogy to spatial probing by coding the mutual intensity, we can modify the temporal coherence function to be $\mathcal{G}(s)$ with both a wavelength-dependent amplitude $A(\kappa)$ and a wavelength-dependent phase shift $\Psi(\kappa)$ on the reference. This performs temporal probing by setting the probing function as:

$$\begin{aligned}\mathcal{P}^c(\epsilon, s) &= \Gamma(\epsilon)\,\mathcal{G}(s)\,. \\ &= \Gamma(\epsilon)\,\mathcal{F}_\kappa\left[S(\kappa)\,A(\kappa)\exp\left(-\mathrm{j}\Psi(\kappa)\right)\right](s)\,. \end{aligned} \tag{6.1}$$

6.2 Passive computational light transport

An exciting future direction is to extend passive interferometry to perform computational light transport. To this end, we combine the interferometric setups of Figure 3.2 and 5.4. We illustrate this combination for sunlight to perform spatial probing in Figure 6.1. As before, we modulate the spatial amplitude and phase of the transverse modes in the interferometer to synthesize convolutional probing functions.

In addition, an exciting future direction is computational interferometric imaging with *any light*. As we mentioned for the result in Figure 5.12, using a near-field light source results breaks the approximation that the transverse modes are plane waves. Correcting for this distortion using wavefront sensing with interferometry can result in, for example, accurate interferometric imaging using passive indoor lighting.

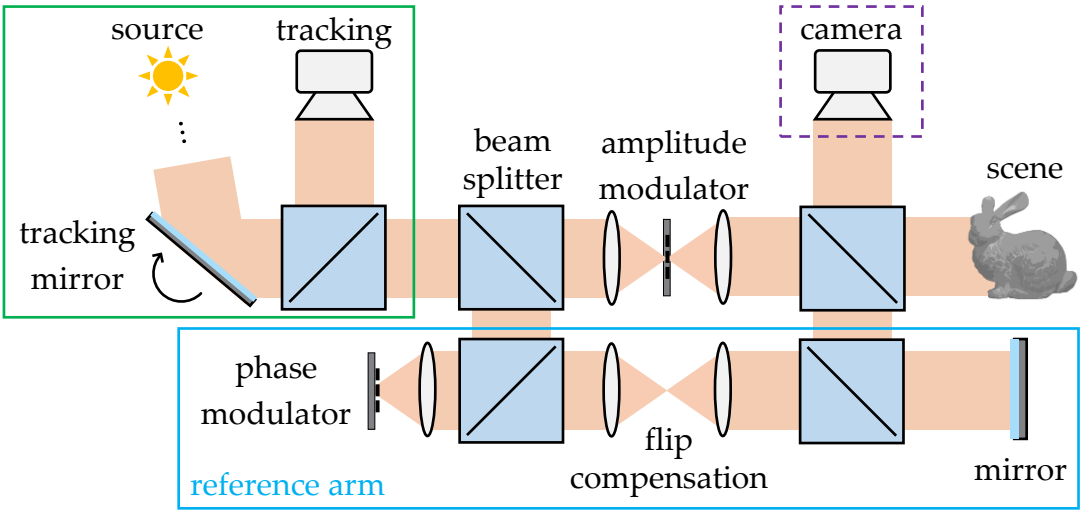


Figure 6.1: **Schematic for proposed passive computational light transport system**.

6.3 Neural processing of interferometric data

Going from raw measurements to desired interference signal in all our interferometry methods (as depicted in Figures 4.5 and 5.3) involves a series of image processing operations: subtraction, squaring, adding, square root, blurring and phase estimation with Equation (4.11). All these operations, along with spatial regularization, are easily incorporated into a deep neural network. We believe that an appropriately designed deep neural network will be able to accomplish end-to-end, high-quality interferometry data processing. This is especially valuable in interferometry with temporally coherent light because of the presence of a huge amount of non-interference speckle noise.

Phase unwrapping in synthetic wavelength interferometry. We saw in Section 4.1 that our method's depth range is limited by the synthetic wavelength. We can attempt to extend our depth range (while not increasing the extent of the reference arm scan) using multiple pairs of lasers to generate multiple different synthetic wavelengths, and using those to perform phase unwrapping. This is a very non-convex problem: In both simulations and preliminary experiments, however, spatially-regularized phase unwrapping algorithms fail to to reconstruct high-quality depth maps in the presence of even slight noise. Such neural data processing can help us perform phase unwrapping to acquire unambiguous depth robustly to noise.

Training such a deep learning pipeline will require collecting depth data for a large variety of scenes. Fortunately, due to the programmable nature of computational interferometric imaging, we can switch between OCT and SWI at the flip of a switch (See

Figure 6.2). This enables us to take SWI measurements as we did in Chapter 4, and very-high-resolution depth (unconstrained by acquisition time) with direct-only OCT (as done by Gkioulekas et al. [18]) for the exact same scene, without the need for spatial registration. We expect that, in addition to phase unwrapping, such a reconstruction pipeline will greatly elevate the quality of the depth results from synthetic wavelength interferometry.

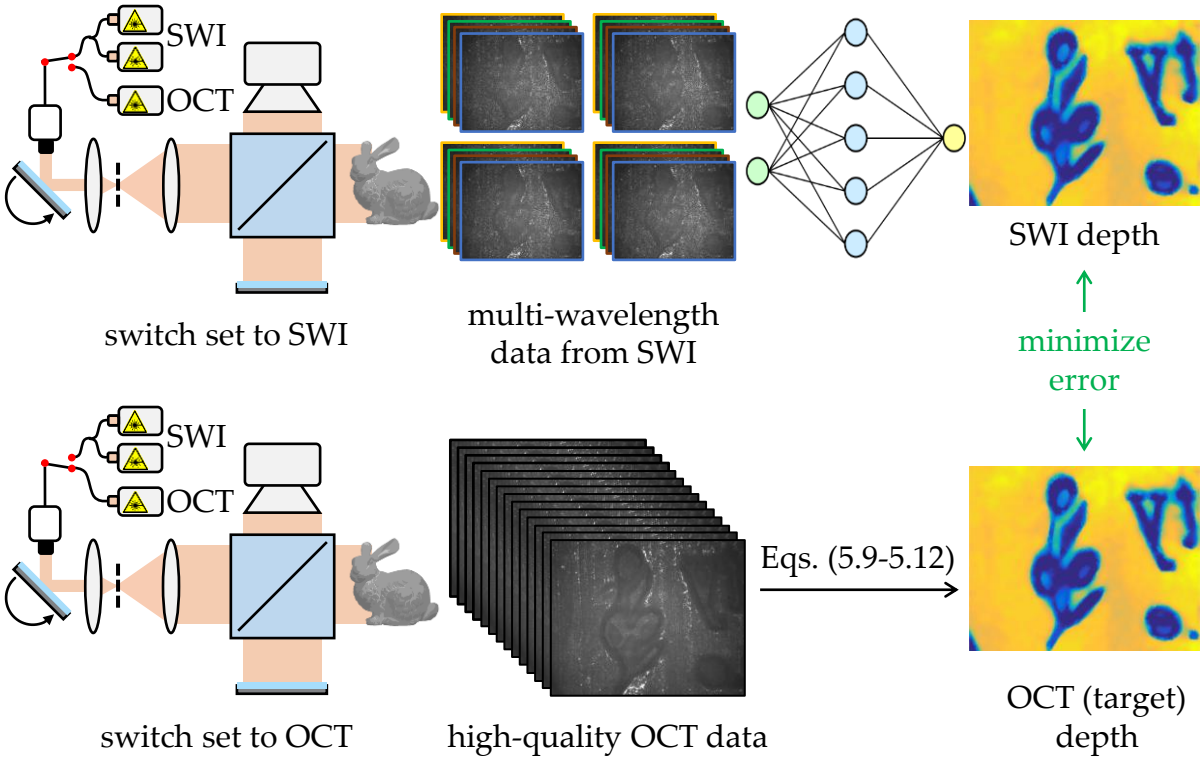


Figure 6.2: **Neural processing of interferometry data example: denoising and phase unwrapping in synthetic wavelength interferometry**.

Appendix A

Interferometry setup for transmission probing

We discuss the engineering details of the setup implementing interferometry with coded mutual intensity. The schematic and a picture of the setup are shown in Figure 3.2.

Light source. We use two types of sources, for temporally coherent and temporally incoherent probing respectively. The first source is a single-frequency green (532 nm) polarized laser from Lasos that outputs a power of 52 mW. This laser has a very small beam diameter (approximately 0.7 mm), requiring that we use of a 5× beam expander to increase the diameter of the beam. The laser is rated to have a coherence length of the order of several meters with a relatively flat temporal coherence function.

The second source is a temporally incoherent supercontinuum laser, *SuperK* from NKT Photonics. The spectrum of this laser ranges from 390 nm to 2400 nm, therefore requiring a spectral filter for a feasible coherence length. We use a 1 nm spectral filter with a central wavelength of 532 nm to reduce the bandwidth of the output light.

Electro-optic modulator. We use an electro-optic modulator to perform amplitude modulation of the beam coming out of the laser. We take advantage of the fact that the beam is thin to pass it through the narrow (2 mm) EOM aperture to pass the beam through. The EOM requires an amplifier converting DC signals of 10 V into DC signals of 200 V. The EOM rotates the polarization of the light according to the signal that is applied to it, and requires the input light to be polarized, and a polarizer on its output. The temporally coherent laser is already polarized, but in the case of the supercontinuum

laser, a polarizer needs to be placed after the EOM. We use a Glan Thompson polarizer on the far end of the EOM and a photographic polarizer in front of the laser if necessary.

Mechanism for spatial incoherence. We use two fast-rotating mirrors to scan the laser beam in a square angle at kilohertz frequencies to create spatial incoherence. An intermediate lens (a 35 mm Nikon prime lens) then maps angle into spatial position behind the illumination lens, creating the 'source'. A beam expander is placed before this entire configuration to expand the very thin beams that the lasers emit, in order to ensure that our final beam covers our field of view of 1 inch. The mirrors are operated by a function generator generating sinusoids at kHz frequency with a slight offset to create a dense Lissajous curve that spans a square. Controlling the size of the source is easy and lossless because that translates to decreasing the amplitude of one of the channels.

Illumination lens. We use a 200 mm Nikon prime lens to collimate light from the above 'source' for its superior performance over off-the-shelf AR-coated achromatic doublets in terms of spherical and chromatic aberration, improving light efficiency and collimation. The output of the scene is cropped to a 1 inch circular beam and is passed through the beamsplitter cube apertures.

Interreflections. Interreflections are especially problematic in the case of temporally coherent light because they introduce strong secondary fringes. For example, light paths passing through a beamsplitter, and those reflected twice inside the beamsplitter also passing through cause strong fringes because of spatial coherence. Such strong fringes in the scene-only and reference-only images essentially nullify our contrast. In the case of temporally incoherent light, this effect is not as severe, because these paths typically have different path lengths and thus do not interfere. We found that deliberately misaligning optics by a small amount (sub-degree) mitigates these effects.

Beamsplitters. We use thin plate 50:50 beamsplitters in all three locations needing beam splitting in our setup. We experimented with pellicle beamsplitters and they introduced strong fringes in the case of coherent light, which reduced our contrast. We did not use cube beamsplitters due to interreflections and the disadvantages discussed above. As above, we deliberately misaligned optics to avoid interreflection artifacts.

We found that in the long path lengths taken by light in our setup, the rough alignment provided by the approximate rigid positioning of the beamsplitters in standard cube mounts is insufficient to get light through the system and into the camera. Therefore, we

used cube mounts that provide control over three degrees-of-freedom of the beamsplitter pose, in order to accuarately set light paths.

Mirrors. We use high-quality mirrors of guaranteed $\lambda/4$ flatness when we do not need phase modulation on the reference arm. In the case of visualizing coherent probing patterns, we use a monolithic hollow mirror retroreflector (three mirrors forming a cubic corner). To create the mirror-diffuser corner in one of the examples, we use a hollow roof mirror and paste a diffuse flat spectral response paper on one of its sides.

Translation stage. We use a translation stage from Newport with an accuracy of upto 10 nm and low-noise operation. Vibrations affect interference contrast, which is why low noise operation is important.

Spatial light modulator lens. To match the lens and magnification of the source on the illumination side, we use another 200 mm Nikon prime lens in front of the phase spatial light modulator.

Phase spatial light modulator. To code the reference beam's phase response, we use a phase spatial light modulator placed in the Fourier plane of the 200 mm Nikon prime lens. The phase spatial light modulator works on the principle of rotating LCD crystals to create a small phase shift on the signal applied to it. However, the input light needs to be polarized in a fixed direction with respect to the SLM to achieve pure phase modulation: we ensure this is the case by setting the polarizer in front of the supercontinuum laser or the orientation of the temporally coherent laser to match the required direction of polarization. For the temporally incoherent case, the entire phase SLM and lens system are placed on the translation stage.

Camera lens. Our scenes are sized at the order of 1 inch. Therefore, we benefit from a lens that achieves high magnifications (1:1). This also allows for better contrast due to lower averaging of speckle (interference signal is convolved with the pixel box when captured with the camera. We use a 180 mm Canon prime macro lens in front of the camera.)

Camera. We use a machine vision camera from AVT with a high sensitivity CCD sensor of resolution 8 MP, pixel size 3.5 μm, and a pixel pitch of 4 μm. Small pixel pitch, along

with pixel size, increase the interference contrast because we average interference speckle over a smaller area.

Color filters. We use a 532 ± 0.2 nm laser line-coated bandpass filter with 1 nm bandwidth and less than 0.01% transmission outside this range, to limit the bandwidth of the supercontinuum laser.

Neutral density filters. We use absorptive neutral density filters to make the intensities of both arms of the interferometer equal. Matching the brightness leads to an acquisition with optimal interference contrast.

Mechanism for inverting the flip due to the lens-and-SLM system. Coding the source with a lens and a phase spatial light modulator behaves as a retroreflector and not a mirror; the flip caused by this arrangement needs to be inverted. We achieve this is to get the reference arm light out of the main beam before the main beamsplitter, have it go inside the phase SLM system and then place a 4f system to flip the beam around before it going to the camera. This motivates the design of the setup in Figure 3.2.

Alignment. Due to the long paths that light takes in the setup and the small (1 inch) aperture, the optical setup requires very careful alignment that can often take hours. Therefore, as much as possible of the setup is built around a cage system that reduces alignment requirements. The optical configuration required to invert the flip, along with the main beamsplitter, can be mounted on a rectangular cage system and clamped down to the optical table for rigidity. The lenses (with the exception of the camera lens) are screwed into the apertures of the cage system with height-adjustable posts below them, and their heights are adjusted to ensure a smooth fit of the lenses in the aperture. Both the lasers are mounted so as to make their beams as parallel to one of the axes of the optical table as possible. The steering mirrors are aligned to make sure the beam at its zero position passes through the center of the beamsplitter box mounts. The beamsplitters are then aligned using a modified version of the alignment technique described by Gkioulekas et al. [18].

Vibrations. The acquisition time of our temporally coherent system is small, because we need to capture either only a single shot or a few subwavelength shifts. Because of the high power of the laser, we need to expose the camera only for a few milliseconds, and therefore vibrations do not strongly affect our measurements. However, when

description	quantity	model name	company
8 MP CCD color camera with Birger EF mount	1	PRO-GT3400-09	Allied Vision Technologies
180 mm compound lens	1	EF 180mm f/3.5L Macro USM	Canon
1 inch laser line bandpass filter, 532 ± 0.2 nm CWL, 1 nm BW	1	FL532-1	Thorlabs
25 mm $\times$ 36 mm plate beamsplitter	3	BSW10R	Thorlabs
motorized filter wheel	1	84-889	Edmund Optics
1 inch round protected Aluminum mirror	3	ME1-G01	Thorlabs
single-frequency laser, 532 nm CWL, 52 mW power	1	unknown	Lasos
broadband hollow retroreflector	1	UBBR2.5-1UV	Newport Corporation
ethernet driver for linear stage	1	XPS-Q2	Newport Corporation
large tilt and rotation platform	2	37	Newport Corporation
ultra-precision linear motor stage, 16 cm travel	1	XMS160	Newport Corporation
200 mm compound lens	2	AF Micro Nikkor 200mm 1:4 D IF-ED	Nikon
phase spatial light modulator	1	Pluto	Holoeye
electro-optic modulator	1	EO-AM-NR-C4	Thorlabs
Glan Thompson coated calcite polarizer	1	GTH5-A	Thorlabs
function generator	2	SDG1025	Siglent
2 inch absorptive neutral density filter kit	1	NEK01S	Thorlabs

Table A.1: **List of major components used in the optical setup of Figure 3.2.**

we combine our system with temporally incoherent light (OCT), the inclusion of the polarizer and spectral filter results in reduced power. Combined with the need to perform a mechanical scan, the acquisition time becomes very long. As a result, in this case, vibrations become problematic, and we need to use a vibration-isolated optical table to conduct our experiments.

Component list. For easy reproducibility of the setup, we provide in Table A.1 a list of the key components used in our implementation. We do not list standard parts used for mounting and positioning, commonly-available in optical labs.

Appendix B

Swept-angle synthetic wavelength interferometry setup

Table B.1: **List of major components used in the optical setup of Figure 4.3(c).**

description	quantity	model name	company
single-frequency lasers, 780 nm CWL, 45 mW power	2	DBR780PN	Thorlabs
benchtop laser diode current controller, ± 500 mA HV	2	LDC205C	Thorlabs
benchtop temperature controller, ± 2 A / 12 WW	2	TED200C	Thorlabs
1 × 2 polarization-maintaining fiber coupler, 780 ± 15 nm	1	PN780R5A1	Thorlabs
reflective FC/APC fiber collimator	1	RC04APC-P01	Thorlabs
2× beam expander	1	GBE02-B	Thorlabs
2-axis galvanometer mirror set	1	GVS202	Thorlabs
function generator	2	SDG1025	Siglent
35 mm compound lens	1	AF Micro Nikkor 35mm 1:4 D IF-ED	Nikon
200 mm compound lens	1	AF Micro Nikkor 200mm 1:4 D IF-ED	Nikon
25 mm × 36 mm plate beamsplitter	3	BSW10R	Thorlabs
1 inch round protected Aluminum mirror	3	ME1-G01	Thorlabs
2 inch absorptive neutral density filter kit	1	NEK01S	Thorlabs
ultra-precision linear motor stage, 16 cm travel	1	XMS160	Newport Corporation
ethernet driver for linear stage	1	XPS-Q2	Newport Corporation
780.5 ± 1 nm OD6 ultra-narrow spectral filter	1	-	Alluxa
180 mm compound lens	1	EF 180mm f/3.5L Macro USM	Canon
8 MP CCD color camera with Birger EF mount	1	PRO-GT3400-09	Allied Vision Technologies

We discuss here the engineering details of the setup implementing swept-angle synthetic wavelength interferometry. The schematic and a picture of the setup are shown in Figure 4.3(c) of the main paper. We use similar components as in the setup of Kotwal et al. [33], and replicate the implementation details below for completeness.

Light source. We use near-infrared single frequency tunable laser diodes from Thorlabs (DBR780PN, 780 nm, 45 mW, 1 MHz linewidth). These laser diodes are tunable in wavelength by adjusting either operating current or temperature of the diode. To create small wavelength separations (of the order of 0.01 nm), we modulate the operating current of

one laser diode with a square waveform, thus create two time-multiplexed wavelengths. To create larger separations (of the order of 1 nm), we use two different laser diodes selected at the appropriate central wavelengths. This is possible because the central wavelengths of separately manufactured laser diodes vary in a ± 2 nm region around 780 nm. We found that for accurate depth recovery, it is important for the light sources to be monochromatic (single longitudinal mode), stable in wavelength and power, and accurately tunable. We experimented with multiple alternatives and encountered problems with either stability, tunability or monochromaticity. We found the DBR lasers from Thorlabs optimal in all these aspects.

Calibrating the synthetic wavelength. The synthetic wavelength resulting from this illumination is very sensitive to the separation between the two wavelengths, especially at microscopic scales. Therefore, after selecting a pair of lasers or current levels for an approximate synthetic wavelength, it is necessary to estimate the actual synthetic wavelength accurately. To do this, we measure the envelope sinusoid for a planar diffuser scene at a dense collection of reference arm positions. We then fit a sinusoid to these measured magnitudes and use the fit wavelength as the synthetic wavelength.

Mechanism for swept-angle scanning. We use two fast-rotating mirrors to scan the laser beam in $1^\circ \times 1^\circ$ angular pattern at slightly separated kHz frequencies, as shown in Chapter 3. A 35 mm Nikon prime lens then maps beam orientation to spatial coordinates behind the illumination lens, creating the light source for swept-angle illumination. The created light source is a dense Lissajous curve that approximates a square. Figure B.1 shows an example light source scanned in this fashion. In practice, we use a much denser light source, but show the coarse one in the inset only to make the Lissajous pattern visible.

Figure B.1: **Lissajous swept-angle source**

Illumination lens. We place the above light source in the focal plane of a 200 mm Nikon prime lens to generate the swept-angle illumination. Photographic lenses perform superior to AR-coated achromatic doublets in terms of spherical and chromatic aberration, therefore resulting in significantly lesser distortion in the generated wavefront.

Interreflections. Interreflections are problematic for us because our illumination is temporally coherent. Interreflections introduce multiple light paths that interfere with each other to create strong spurious fringes. Such fringes suppress the contrast of our speckle signal. The optics we use are coated with anti-reflective films designed for our laser wavelengths to reduce interreflections. We also deliberately misalign our optics with sub-degree rotations from the ideal alignment to avoid strong interreflections.

Beamsplitter. We use a thin 50:50 plate beamsplitter, since pellicle and cube beamsplitters cause strong fringes. As above, we deliberately misalign the beamsplitter to avoid interreflections.

Mirrors. We use high-quality mirrors of guaranteed $\lambda/4$ flatness to ensure a uniform phase reference throughout the field of view of the camera.

Translation stage. We use a ultra-precision motorized linear translation stage from Newport with a positioning accuracy of up to 10 nm and minimum incremental motion of 1 nm. For high-resolution depth recovery, it is important that the mirror positions images are captured at be accurate at sub-optical wavelength scales. In addition, it is important that the translation stage guarantee low-positioning-noise operation.

Camera lens. Our scenes are sized at the order of 1 inch. Therefore, we benefit from a lens that achieves high magnifications (1:1 reproduction ratio). This also allows for better contrast due to lower averaging of speckle (interference signal is convolved with the pixel box when captured with the camera). We use a 180 mm Canon prime macro lens in front of the camera.

Camera. We use a machine vision camera from Allied Vision with a high-sensitivity CCD sensor of 8 MP resolution, and pixel size 3.5 µm. A sensor with a small pixel pitch averages interference speckle over a smaller spatial area, therefore allowing us to resolve finer lateral detail. We use a camera with the protective glass of sensor removed. This is critical to avoid spurious interreflections caused by the protective glass.

Neutral density filters. We use absorptive neutral density filters to optimize interference contrast by making the intensities of both interferometer arms equal.

Alignment. For depth estimation accurate to micron-scales, the optical setup requires very careful alignment. To avoid as much human error as possible, we build the illumination side and the beamsplitter holder on a rigid cage system constructed with components from Thorlabs. To ensure a mean direction of light propagation that's parallel to the optical axis of the interferometer, we tune the steering mirrors electronically by adjusting their driving waveform's DC offset. We then align the reference mirror and camera using the alignment technique described by Gkioulekas et al. [18].

Component list. For reproducibility, Table B.1 gives a list of the important components used in our implementation.

Appendix C

Passive sunlight interferometry setup

We discuss implementation details of our passive interferometry setup in Figure 5.4, and provide a full parts list in Table C.1. We take inspiration from the interferometer designs by Gkioulekas et al. [18] and Kotwal et al. [33], and modify them to use sunlight and achieve robustness to environmental conditions to achieve high-quality depth sensing outdoors. Wherever it is necessary to align optical components, we use the methods described by Gkioulekas et al. [18].

Vibration isolation. The factor most detrimental to the quality of our results is vibrations in uncontrolled outdoor conditions. Vibrations arise from multiple sources: (a) mechanical vibrations from the ground that propagate up the cart because it has no suspension, (b) movement of the optical components on the table with respect to each other and (c) lateral and axial movement of the scene and reference and tracking mirrors induced by strong winds.

To mitigate (a), we place our setup on a passively damped honeycomb optical breadboard from Thorlabs on the upper level of our utility cart. We found in experiments that using the breadboard and eliminating any other sources of vibrations (such as the computer in Figure 5.4(b) and cables connected to the various components) on the upper level of the cart is sufficient. We tightly route the cables so they do not move parts of the setup. With the damped breadboard, vibrations from the lower level of the cart do not travel to the optical setup, so any noisy equipment such as the reference arm translation stage controller (the white cabinet in Figure 5.4(b)) can be placed on the lower level.

To mitigate (b), we modified the interferometer design from Kotwal et al. [33] so as to tightly clamp down the optical components to each other and to the breadboard.

We achieved this by connecting the beamsplitters and the cameras with a 30 mm cage system using components from Thorlabs. It is not practical to connect the tracking and reference mirrors similarly, so we clamped them tightly to the breadboard using mounting components from Thorlabs.

The movement of the mirrors in (c) can be mitigated by choosing high-resistance translation stages and high-torque rotation stages. We comment on the properties of our chosen motion stages below. We found that our setup is able to tolerate winds as fast as 25 mph.

Tracking the sun. As we mention in the main paper, it is important to track the sun and center the sunlight beam along the optical axis of the system several times during acquisition. To keep the distribution of sunlight across the field of view uniform, we found that it is necessary to re-center the beam every 100 intensity images. In the duration of capturing 100 images at 50 ms per image, the sun moves 0.02 $^\circ$ in the sky. We perform the tracking using two motorized precision rotation stages from Thorlabs, one for the azimuth and one for the altitude of the mirror. These stages have minimum incremental motions of 0.03 $^\circ$, making them perfect for our setup. We found in tests outdoors that the torque provided by these rotation stages was enough to keep the tracking mirror stable in winds as fast as 25 mph. As the control system, we found that a proportional controller with k_p set as half the sensitivity of pixel displacements to rotation stage angles was sufficient.

Reference translation stage. To ensure that we place the reference mirror at the desired positions accurately enough for micron-scale resolutions, we need a translation stage that has a minimum incremental translation less than one micron. We use the XMS160 translation stage from Newport that has a minimal incremental translation of 10 nm. In addition, this stage guarantees low-noise and high-resistance operation, preventing loss of interference contrast due to reference mirror position noise caused by mechanical vibrations and flowing wind.

Beamsplitters. We use thin plate beamsplitters from Thorlabs. The beamsplitter sending light to the tracking camera is a 10:90 (R:T) beamsplitter, reflecting 10% of the input light to the tracking camera. The imaging beamsplitter is a 50:50 (R:T) beamsplitter. We choose plate beamsplitters over cube beamsplitters because cubes cause significant interreflections, and over pellicle beamsplitters because of their tendency to distort due to airflow.

Table C.1: **List of major components used in the optical setup in Figure 5.4(b).**

description	quantity	model name	company
3′ × 2′ utility cart	1	https://www.amazon.com/dp/B001602VI2	Rubbermaid
passively damped optical breadboard	1	B2436FX	Thorlabs
2.56″ motorized precision rotation stage	1	PRMTZ8	Thorlabs
1″ motorized precision rotation stage	1	PRM1Z8	Thorlabs
K-Cube brushed DC servo motor controller	2	KDC101	Thorlabs
2″ round protected aluminum mirror	1	ME2-G01	Thorlabs
25 × 36 mm plate beamsplitter, 10:90 (R:T)	1	BSN10R	Thorlabs
25 × 36 mm plate beamsplitter, 50:50 (R:T)	1	BSW10R	Thorlabs
300 mm compound lens	2	AF Micro Nikkor 300mm 1:4 D IF-ED	Nikon
tracking camera	1	Grasshopper3 USB3 GS3-U3-41C6M-C	FLIR
imaging camera	1	Blackfly S USB3 BFS-U3-122S6M-C	FLIR
1″ round protected aluminum mirror	1	ME1-G01	Thorlabs
2″ absorptive neutral density filter kit	1	NEK03	Thorlabs
ultra-precision linear motor stage, 16 cm travel	1	XMS160	Newport Corporation
ethernet driver for linear stage	1	XPS-Q2	Newport Corporation
550 ± 20 nm bandpass spectral filter	1	FB550-40	Thorlabs

Mirrors. We use high-quality protected aluminum mirrors of guaranteed $\lambda/4$ flatness to ensure a uniform phase distribution throughout the sunlight beam.

Spectral filter. As mentioned in the main paper, we control the spectral bandwidth of sunlight using a spectral filter. We choose the spectral filter to balance light efficiency, acquisition time, temporal coherence length, and signal to noise ratio. If we increase the spectral filter bandwidth, the acquisition time decreases, the signal to ratio increases, and the temporal coherence length decreases. The latter makes depth recovery more sensitive to vibrations from the environment. We found in experiments that a spectral filter with central wavelength 550 nm and bandwidth 20 nm works best with out setup.

Alignment and calibration. In addition to handling the factors detrimental to interferometric depth reconstruction that we detailed above, we need to deal with the imperfections within the system's construction. The stability that the cage system provides comes at the loss of the capability to finely adjust the orientations and positions of individual components. We need to calibrate for the errors caused by misalignment. In addition, the tracking and control system parameters are critical to set properly for sun tracking to work fast enough without becoming unstable. With this in view, there are eleven important calibration parameters in the setup for accurate light delivery and tracking stability.

1. Tracking camera principal point: Due to small misalignments in the cage system, when the tracking mirror deflects light perfectly parallel to the optical axis of the system, the image of the sun at infinity does not appear at the center of the tracking damera sensor. We manually adjust the tracking mirrors to make sure that

the shadows of the cage system rods on the beamsplitter mounts disappear. The position of center of the sun image at this tracking mirror position is the principal point of the tracking camera.

2. Tracking motor speeds and accelerations: The tracking mirror rotation stages have controllable limits on angular speeds and accelerations. We tune these to achieve the minimum incremental motion at the principal point as fast as possible while still being stable.
3. Sensitivity of pixel displacements on the camera to tracking mirror angular displacements: To get a ballpark number for the proportional controller gain, we estimate this sensitivity by rotating the tracking mirror in both dimensions by the minimum incremental rotation of the tracking mirror stages and measure the displacement of the sun image center.
4. Proportional controller gains: Finally, the proportional controllers we use to output control signals to the rotation stages need to be as fast as possible without overshooting. Overshooting can be catastrophic, because a large overshoot leads the sun out of the field of view of the tracking camera.
5. Time between tracking adjustments: The movement of the sun causes spatial intensity non-uniformities in the captured scan. To minimize this while considering the minimum incremental motion of the rotation stages, we calibrate how many scan images to capture between adjustments using the exposure time per image.

Bibliography

[1] Supreeth Achar, Joseph R Bartels, William L Whittaker, Kiriakos N Kutulakos, and Srinivasa G Narasimhan. Epipolar time-of-flight imaging. *ACM TOG*, 2017. 3.7

[2] Girish S. Agarwal, Greg Gbur, and Emil Wolf. Coherence properties of sunlight. *Opt. Lett.*, 29(5):459–461, Mar 2004. doi: 10.1364/OL.29.000459. URL http://opg.optica.org/ol/abstract.cfm?URI=ol-29-5-459. 5.4

[3] Seung-Hwan Baek, Noah Walsh, Ilya Chugunov, Zheng Shi, and Felix Heide. Centimeter-wave free-space neural time-of-flight imaging. *ACM TOG*, 2022. 4.5

[4] Jacopo Bertolotti, Elbert G Van Putten, Christian Blum, Ad Lagendijk, Willem L Vos, and Allard P Mosk. Non-invasive imaging through opaque scattering layers. *Nature*, 2012. 1.1, 3

[5] Max Born, Emil Wolf, A. B. Bhatia, P. C. Clemmow, D. Gabor, A. R. Stokes, A. M. Taylor, P. A. Wayman, and W. L. Wilcock. *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light*. Cambridge University Press, 7 edition, 1999. doi: 10.1017/CBO9781139644181. 2.1

[6] S. Chandrasekhar. *Radiative transfer*. Dover, 1960. 1.3

[7] Chengqian Che, Fujun Luan, Shuang Zhao, Kavita Bala, and Ioannis Gkioulekas. Towards learning-based inverse subsurface scattering. In *2020 IEEE International Conference on Computational Photography (ICCP)*, pages 1–12. IEEE, 2020. 2

[8] Yeou-Yen Cheng and James C. Wyant. Two-wavelength phase shifting interferometry. *Applied Optics*, 1984. 4.5

[9] Yeou-Yen Cheng and James C. Wyant. Multiple-wavelength phase-shifting interferometry. *Applied Optics*, 1985. 4.5

[10] Oliver Cossairt, Nathan Matsuda, and Mohit Gupta. Digital refocusing with incoherent holography. *IEEE ICCP*, 2014. 1.3, 5.3

[11] Katherine Creath. Phase-shifting speckle interferometry. *Applied Optics*, 1985. 3.5

[12] Peter de Groot. *Phase Shifting Interferometry*. 2011. 1.3, 2.3, 4.3, 5.2, 5.2

[13] David Droeschel, Dirk Holz, and Sven Behnke. Multi-frequency phase unwrapping for time-of-flight cameras. *IEEE/RSJ IROS*, 2010. 4.5

[14] Elmar Eisemann and Frédo Durand. Flash photography enhancement via intrinsic relighting. *ACM TOG*, 2004. 4.3

[15] Adolf F Fercher, Wolfgang Drexler, Christoph K Hitzenberger, and Theo Lasser. Optical coherence tomography-principles and applications. *Reports on progress in physics*, 2003. 5.1

[16] James R Fienup. Phase retrieval algorithms: a comparison. *Applied optics*, 1982. 3.7

[17] Genevieve Gariepy, Nikola Krstajić, Robert Henderson, Chunyong Li, Robert R Thomson, Gerald S Buller, Barmak Heshmat, Ramesh Raskar, Jonathan Leach, and Daniele Faccio. Single-photon sensitive light-in-fight imaging. *Nature Comm.*, 2015. 1

[18] Ioannis Gkioulekas, Anat Levin, Frédo Durand, and Todd Zickler. Micron-scale light transport decomposition using interferometry. *ACM TOG*, 2015. 1.1, 1.3, 2.3, 3.5, 3.6, 3.7, 3.7, 4.3, 4.4, 5.1, 5.2, 5.3, 5.3, 6.3, A, B, C

[19] Ioannis Gkioulekas, Anat Levin, and Todd Zickler. An evaluation of computational imaging techniques for heterogeneous inverse scattering. *ECCV*, 2016. 1.1, 3.6, 1

[20] Joseph W Goodman. *Introduction to Fourier optics*. Roberts and Company Publishers, 2005. 2.1, 2.1

[21] Mohit Gupta, Shree K Nayar, Matthias B Hullin, and Jaime Martin. Phasor imaging: A generalization of correlation-based time-of-flight imaging. *ACM TOG*, 2015. 3.5, 4.5

[22] Parameswaran Hariharan. *Optical interferometry*. Elsevier, 2003. 1.3, 1.3

[23] Samuel W Hasinoff, Frédo Durand, and William T Freeman. Noise-optimal capture for high dynamic range photography. *CVPR*, 2010. 3.7

[24] Felix Heide, Matthias B. Hullin, James Gregson, and Wolfgang Heidrich. Low-budget Transient Imaging Using Photonic Mixer Devices. *ACM Trans. Graph.*, 32(4): 45:1–45:10, July 2013. ISSN 0730-0301. 1

[25] David Huang, Eric A Swanson, Charles P Lin, Joel S Schuman, William G Stinson, Warren Chang, Michael R Hee, Thomas Flotte, Kenton Gregory, Carmen A Puliafito, and James G Fujimoto. Optical coherence tomography. *Science*, 1991. 1.3, 2.3

[26] Benjamin Judkewitz, Roarke Horstmeyer, Ivo M Vellekoop, Ioannis N Papadopoulos, and Changhuei Yang. Translation correlations in anisotropically scattering media. *Nature physics*, 2015. 3.2

[27] Achuta Kadambi, Refael Whyte, Ayush Bhandari, Lee Streeter, Christopher Barsi, Adrian Dorrington, and Ramesh Raskar. Coded Time of Flight Cameras: Sparse Deconvolution to Address Multipath Interference and Recover Time Profiles. *ACM TOG*, 2013. 1

[28] Takashi Kakue, Ryosuke Yonesaka, Tatsuki Tahara, Yasuhiro Awatsuji, Kenzo Nishio, Shogo Ura, Toshihiro Kubota, and Osamu Matoba. High-speed phase imaging by parallel phase-shifting digital holography. *Optics letters*, 2011. 3.7

[29] Ori Katz, Pierre Heidmann, Mathias Fink, and Sylvain Gigan. Non-invasive single-shot imaging through scattering layers and around corners via speckle correlations. *Nature photonics*, 2014. 3

[30] Moonseok Kim, Wonjun Choi, Youngwoon Choi, Changhyeong Yoon, and Wonshik Choi. Transmission matrix of a scattering medium and its applications in biophotonics. *Opt. Express*, 23(10):12648–12668, May 2015. doi: 10.1364/OE.23.012648. URL https://opg.optica.org/oe/abstract.cfm?URI=oe-23-10-12648. 1.1

[31] Ahmed Kirmani, Tyler Hutchison, James Davis, and Ramesh Raskar. Looking around the corner using ultrafast transient imaging. *IJCV*, 2011. 1

[32] Johannes Kopf, Michael F. Cohen, Dani Lischinski, and Matt Uyttendaele. Joint bilateral upsampling. *ACM Trans. Graph.*, 26(3):96–es, jul 2007. ISSN 0730-0301. doi: 10.1145/1276377.1276497. URL https://doi.org/10.1145/1276377.1276497. 4.4

[33] Alankar Kotwal, Anat Levin, and Ioannis Gkioulekas. Interferometric transmission probing with coded mutual intensity. *ACM TOG*, 2020. 4.3, 5.3, B, C, C

[34] Alankar Kotwal, Anat Levin, and Ioannis Gkioulekas. Project website, 2020. https://imaging.cs.cmu.edu/coded_mutual_intensity. 3.6, 3.6, 3.6, 3.6, 3.7

[35] Hiroyuki Kubo, Suren Jayasuriya, Takafumi Iwaguchi, Takuya Funatomi, Yasuhiro Mukaigawa, and Srinivasa G Narasimhan. Acquiring and characterizing plane-to-ray indirect light transport. *IEEE ICCP*, 2018. 1, 3.6

[36] Guanming Lai and Toyohiko Yatagai. Generalized phase-shifting interferometry. *JOSA A*, 1991. 4.3

[37] Szu-Yu Lee, Vicente J. Parot, Brett E. Bouma, and Martin Villiger. Confocal 3d reflectance imaging through multimode fiber without wavefront shaping. 2

[38] Fengqiang Li, Joshua Yablon, Andreas Velten, Mohit Gupta, and Oliver Cossairt. High-depth-resolution range imaging with multiple-wavelength superheterodyne interferometry using 1550-nm lasers. *Applied Optics*, 2017. 1.3, 5.3

[39] Fengqiang Li, Florian Willomitzer, Prasanna Rangarajan, Mohit Gupta, Andreas Velten, and Oliver Cossairt. Sh-tof: Micro resolution time-of-flight imaging with superheterodyne interferometry. *IEEE ICCP*, 2018. 1.3, 5.3

[40] Chao Liu, Akash Maity, Artur W Dubrawski, and Srinivasa G Sabharwal, Ashutosh Narasimhan. High resolution diffuse optical tomography using short range indirect subsurface imaging. *IEEE ICCP*, 2020. 1, 3.6

[41] Chao Liu, Akash K. Maity, Artur W. Dubrawski, Ashutosh Sabharwal, and Srinivasa G. Narasimhan. High resolution diffuse optical tomography using short range indirect subsurface imaging. In *2020 IEEE International Conference on Computational Photography (ICCP)*, pages 1–12, 2020. doi: 10.1109/ICCP48838.2020.9105173. 1.1

[42] Tomohiro Maeda, Achuta Kadambi, Yoav Y Schechner, and Ramesh Raskar. Dynamic heterodyne interferometry. *IEEE ICCP*, 2018. 5.3

[43] Heylal Mashaal, Alex Goldstein, Daniel Feuermann, and Jeffrey M. Gordon. First direct measurement of the spatial coherence of sunlight. *Opt. Lett.*, 37(17):3516–3518, Sep 2012. doi: 10.1364/OL.37.003516. URL http://opg.optica.org/ol/abstract.cfm?URI=ol-37-17-3516. 5.4

[44] Christopher A Metzler, Manoj K Sharma, Sudarshan Nagesh, Richard G Baraniuk, Oliver Cossairt, and Ashok Veeraraghavan. Coherent inverse scattering via transmission matrices: Efficient phase retrieval algorithms and a public dataset. *IEEE ICCP*, 2017. 1.1, 3

[45] Eman Mirdamadi, Joshua W. Tashman, Daniel J. Shiwarski, Rachelle N. Palchesko, and Adam W. Feinberg. Fresh 3d bioprinting a full-size model of the human heart. *ACS Biomaterials Science & Engineering*, 6(11):6453–6459, Nov 2020. doi: 10.1021/acsbiomaterials.0c01133. URL https://doi.org/10.1021/acsbiomaterials.0c01133. 4

[46] Inc. Mirrorcle Technologies. Mirrorcle mems technical overview, 2022. https://www.mirrorcletech.com/pdf/Mirrorcle_MEMS_Mirrors_-_Technical_Overview.pdf. 2, 3

[47] Allard P. Mosk, Ad Lagendijk, Geoffroy Lerosey, and Mathias Fink. Controlling waves in space and time for imaging and focusing in complex media. *Nature*

Photonics, 6(5):283–292, May 2012. ISSN 1749-4893. doi: 10.1038/nphoton.2012.88. URL https://doi.org/10.1038/nphoton.2012.88. 1.1

[48] Matthew O'Toole. *Optical Linear Algebra for Computational Light Transport*. PhD thesis, University of Toronto (Canada), 2016. 3.7

[49] Matthew O'Toole and Kiriakos N Kutulakos. Optical Computing for Fast Light Transport Analysis. *ACM TOG*, 2010. 2, 1.3, 3

[50] Matthew O'Toole, Ramesh Raskar, and Kiriakos N Kutulakos. Primal-dual Coding to Probe Light Transport. *ACM TOG*, 2012. 1, 3, 3.2, 1, 3.6, 3.6, 3.6, 3.7

[51] Matthew O'Toole, Felix Heide, Lei Xiao, Matthias B Hullin, Wolfgang Heidrich, and Kiriakos N Kutulakos. Temporal Frequency Probing for 5D Transient Analysis of Global Light Transport. *ACM TOG*, 2014. 1.1, 3, 3.2, 3.7, 5.1

[52] Matthew O'Toole, John Mather, and Kiriakos N Kutulakos. 3d shape and indirect appearance by structured light transport. *IEEE/CVF CVPR*, 2014. 1, 3.7

[53] Matthew O'Toole, Supreeth Achar, Srinivasa G Narasimhan, and Kiriakos N Kutulakos. Homogeneous codes for energy-efficient illumination and imaging. *ACM TOG*, 2015. 1, 3.4, 3.7

[54] Matthew O'Toole, Felix Heide, David B Lindell, Kai Zang, Steven Diamond, and Gordon Wetzstein. Reconstructing transient images from single-photon sensors. *IEEE/CVF CVPR*, 2017. 1

[55] Christoph Peters, Jonathan Klein, Matthias B Hullin, and Reinhard Klein. Solving trigonometric moment problems for fast transient imaging. *ACM Trans. Graph.*, 34 (6):220–1, 2015. 1

[56] Georg Petschnigg, Richard Szeliski, Maneesh Agrawala, Michael Cohen, Hugues Hoppe, and Kentaro Toyama. Digital photography with flash and no-flash image pairs. *ACM TOG*, 2004. 4.3

[57] Sébastien Popoff, Geoffroy Lerosey, Mathias Fink, Albert Claude Boccara, and Sylvain Gigan. Image transmission through an opaque material. *Nature communications*, 1(1):1–5, 2010. 1.1, 3

[58] SM Popoff, G Lerosey, R Carminati, M Fink, AC Boccara, and S Gigan. Measuring the transmission matrix in optics: an approach to the study and control of light propagation in disordered media. *Physical review letters*, 104(10):100601, 2010. 1.1, 3

[59] Berke Vow Ricketti, Erik M Gauger, and Alessandro Fedrizzi. The coherence time of sunlight in the context of natural and artificial light-harvesting. *Scientific Reports*, 12

(1):1–9, 2022. 5.1

[60] Stefan Rothe, Hannes Radner, Nektarios Koukourakis, and Jürgen W. Czarske. Transmission matrix measurement of multimode optical fibers by mode-selective excitation using one spatial light modulator. *Applied Sciences*, 9(1), 2019. ISSN 2076-3417. doi: 10.3390/app9010195. URL https://www.mdpi.com/2076-3417/9/1/195. 2

[61] Haowen Ruan, Joshua Brake, J Elliott Robinson, Yan Liu, Mooseok Jang, Cheng Xiao, Chunyi Zhou, Viviana Gradinaru, and Changhuei Yang. Deep tissue optical focusing and optogenetic modulation with time-reversed ultrasonically encoded light. *Sci Adv*, 3(12):eaao5520, December 2017. 2

[62] Manoj Sharma, Christopher A Metzler, Sudarshan Nagesh, Oliver Cossairt, Richard G Baraniuk, and Ashok Veeraraghavan. Inverse scattering via transmission matrices: Broadband illumination and fast phase retrieval algorithms. *IEEE TCI*, 2019. 1.1, 3

[63] Pitchaya Sitthi-Amorn, Javier E Ramos, Yuwang Wangy, Joyce Kwan, Justin Lan, Wenshou Wang, and Wojciech Matusik. Multifab: a machine vision assisted platform for multi-material 3d printing. *ACM TOG*, 34(4):129, 2015. 4

[64] Orazio Svelto. *Principles of Lasers*. 2010. 1

[65] Kazumasa Takada. Noise in optical low-coherence reflectometry. *IEEE Journal of Quantum Electronics*, 1998. 3.7

[66] I. M. Vellekoop and A. P. Mosk. Universal optimal transmission of light through disordered materials. *Phys. Rev. Lett.*, 101:120601, Sep 2008. doi: 10.1103/PhysRevLett.101.120601. URL https://link.aps.org/doi/10.1103/PhysRevLett.101.120601. 1.1

[67] Ivo M Vellekoop and AP Mosk. Focusing coherent light through opaque strongly scattering media. *Optics letters*, 2007. 1.1, 3

[68] Ivo M Vellekoop, Meng Cui, and Changhuei Yang. Digital optical phase conjugation of fluorescence in turbid tissue. *Applied Physics Letters*, 2012. 3

[69] Andreas Velten, Di Wu, Adrian Jarabo, Belen Masia, Christopher Barsi, Chinmaya Joshi, Everett Lawson, Moungi Bawendi, Diego Gutierrez, and Ramesh Raskar. Femto-photography: capturing and visualizing the propagation of light. *ACM TOG*, 2013. 1

[70] Jian Wang, Joseph Bartels, William Whittaker, Aswin C Sankaranarayanan, and

Srinivasa G Narasimhan. Programmable triangulation light curtains. *ECCV*, 2018. 1

[71] Di Wu, Andreas Velten, Matthew O'Toole, Belen Masia, Amit Agrawal, Qionghai Dai, and Ramesh Raskar. Decomposing Global Light Transport Using Time of Flight Imaging. *IJCV*, 2014. 1

[72] Di Wu, Gordon Wetzstein, Christopher Barsi, Thomas Willwacher, Qionghai Dai, and Ramesh Raskar. Ultra-fast Lensless Computational Imaging through 5D Frequency Analysis of Time-resolved Light Transport. *IJCV*, 2014. 1

[73] Shumian Xin, Sotiris Nousias, Kiriakos N. Kutulakos, Aswin C. Sankaranarayanan, Srinivasa G. Narasimhan, and Ioannis Gkioulekas. A theory of fermat paths for non-line-of-sight shape reconstruction. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6793–6802, 2019. doi: 10.1109/CVPR.2019.00696. 1.3