Micron-scale Light Transport Decomposition Using Interferometry: Supplementary Material

Ioannis Gkioulekas, Harvard University
Anat Levin, Weizmann Institute
Frédo Durand, Massachusetts Institute of Technology
Todd Zickler, Harvard University

Abstract

We present a computational imaging system, inspired by the optical coherence tomography (OCT) framework, that uses interferometry to produce decompositions of light transport in small scenes or volumes. The system decomposes transport according to various attributes of the paths that photons travel through the scene, including where on the source the paths originate, their pathlengths from source to camera through the scene, their wavelength, and their polarization. Since it uses interference, the system can achieve high pathlength resolutions, with the ability to distinguish paths whose lengths differ by as little as ten microns. We describe how to construct and optimize an optical assembly for this technique, and we build a prototype to measure and visualize three-dimensional shape, direct and indirect reflection components, and properties of scattering, refractive/dispersive, and birefringent materials.

CR Categories: I.4.1 [Image Processing and Computer Vision]: Digitization and Image Capture—Imaging geometry;

Keywords: light transport, interference, wave optics

1 Introduction

In imaging, photons leave a source, travel through a scene, and are collected by a camera. A conventional image measures the sum of all photons that arrive at each camera pixel, regardless of where on the source they originate, or which path they travel from there to the camera (Figure 1). This summation conflates information about the shapes and materials in the scene. In this paper, we show how interferometry can be used to measure decompositions of these per-pixel sums, distinguishing between photon paths that differ in their length and location of origin.

In discrete terms, where the camera sensor is partitioned into P area elements ("pixels") and the source is partitioned into L area elements, the energy that is transferred from source to camera is often described by a $P \times L$ matrix $\mathbf{T}$ called the *light transport matrix*. Each entry T_{pl} in this matrix represents the fraction of photons following paths originating at the lth source element and arriving at the pth camera element. With reference to Figure 1, orange versus blue paths would correspond to entries T_{pl} for different values of source location l. Conventional imaging simply measures the image $\mathbf{i}$ (a P-vector) for some pattern $\boldsymbol{l}$ (an L-vector) projected from the source according to the *light transport equation* [Ng et al. 2003]

$$\mathbf{i} = \mathbf{T}\boldsymbol{l}. \tag{1}$$

There are a variety of computational imaging methods for measuring elements of the transport matrix. This has applications in image-based rendering, image editing, and measuring scene shape in the presence of translucency and interreflections.

Each element of the transport matrix in Equation (1) still represents a sum of different scene paths. As depicted by the blue paths of Figure 1, sub-surface scattering and interreflections mean that there are typically many paths that originate at the same source element and arrive at the same camera element but take different routes through the scene. Computational imaging can be used to analyze these as well, by decomposing each entry T_{pl} of the transport matrix according to the contributions from photon paths that have different optical lengths τ. This leads to the notion of a *pathlength-resolved light transport matrix* $\mathbf{T}^\tau$, $\tau \in \{\tau_{\text{min}}, \ldots, \tau_{\text{max}}\}$, where pathlengths are discretized into a finite set of $\Delta\tau$-sized length intervals. The full light transport matrix can then be written as

$$\mathbf{T} = \sum_\tau \mathbf{T}^\tau. \tag{2}$$

Each entry T^τ_{pl} is the fraction of photons originating at the lth source element that arrive at the pth camera element having traveled paths whose optical lengths are in the interval $\tau \pm \Delta\tau/2$. There is a growing number of methods for measuring either entire projections or isolated elements of the pathlength-resolved transport matrix. In addition to enhancing the applications listed above, these decompositions can be used to visualize light-in-flight through table-top scenes, and for "imaging around corners."

We introduce a new computational imaging system that uses optical interferometry to produce high-fidelity light transport decompositions. Our system uses optical configurations that are variations of the classical Michelson interferometer, and our analysis builds on techniques that have been used for optical coherence tomography (OCT). Our system is complementary to existing computational photography methods for producing such decompositions, excelling in use cases where it is necessary to image small scenes at very high spatial and pathlength resolutions. In particular, the use of interferometry allows our system to achieve pathlength resolutions as low as $10\,\mu\text{m}$, which is necessary to analyze transport events caused by material effects like dispersion and scattering.

Our paper begins with background on the Michelson interferometer and the notions of spatial and temporal coherence length. We then introduce a mathematical model of interferometry in terms of (a continuous version of) the pathlength-resolved light transport matrix. We use it to show how sources with different coherence properties enable different kinds of light transport decompositions, differentiating light paths in terms of their endpoint locations, their optical lengths, or combinations of these two. We also characterize resolution and noise performance, and present a performance-optimized optical design that additionally allows resolving transport in terms of wavelength and polarization. Our prototype has three spectral channels, two polarization channels, a working volume of $2\,\text{cm}$ H $\times$ $2\,\text{cm}$ W $\times$ $1\,\text{cm}$ D, and spatial and pathlength resolutions that are both $10\,\mu\text{m}$. We use this prototype to obtain micron-scale decompositions of light transport in scenes containing reflection, refraction, dispersion, scattering, and birefringence.

2 Related Work

Light transport decomposition. Methods for decomposing light transport, both spatially and in terms of pathlength, can be described as capturing different images $\mathbf{i}$ of a static scene under uni-

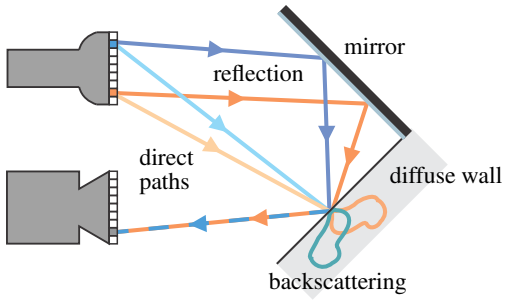


Figure 1: *Differentiating light paths. A pixel in a conventional image measures the sum of all lights paths arriving at that pixel. Entries of a light transport matrix distinguish paths with different starting locations (orange vs. blue), and entries of a "pathlength-resolved" light transport matrix additionally distinguish paths of different lengths (e.g., light vs. dark blue).*

form illumination $\boldsymbol{l} = \mathbf{1}$ according to the equation

$$\mathbf{i} = \sum_{\tau} w(\tau)(\mathbf{M} \odot \mathbf{T}^{\tau})\mathbf{1}. \tag{3}$$

This is a pathlength-resolved variant of the transport probing equation of O'Toole et al. [2012]. The operator $\odot$ represents pointwise multiplication (Hadamard product). $\mathbf{M}$ is a $P \times L$ binary matrix: setting M_{pl} to zero removes contributions of paths beginning at point l on the source and arriving at pixel p on the sensor. Similarly, $w(\tau)$ is a scalar binary function that can be used to remove contributions from paths of length τ. In conventional imaging, $\mathbf{M}$ and $w(\tau)$ are equal to 1 everywhere, and Equation (3) reduces to the standard light transport Equation (1) with uniform illumination. Computational imaging methods for light transport decomposition collect measurements with different $\mathbf{M}$ and $w(\tau)$. We will use Figure 1 to visualize paths captured by various approaches.

Spatial decomposition methods induce different matrices $\mathbf{M}$ while keeping $w(\tau) = 1$ everywhere. With reference to Figure 1, these methods can be used to record contributions from only the blue or only the orange paths. By using different arrangements of a camera and a projector, and different sensor and illumination coding strategies, these methods capture individual elements of the transport matrix, or sparse or low-rank approximations to it [Sen et al. 2005; Peers et al. 2009; Wang et al. 2009; Bai et al. 2010; O'Toole and Kutulakos 2010; O'Toole et al. 2012; O'Toole et al. 2014a].

There are also methods that decompose according to pathlength only, using a matrix $\mathbf{M} = \mathbf{1}$ and controllable functions $w(\tau)$. Referring to Figure 1, this allows measuring the superposition of all paths (blue and orange) that have the same length. This has been dubbed *transient imaging* and can be used to produce light-in-flight visualizations of photons propagating in a scene. Implementations include using a combination of pulsed laser and ultra-fast detector [Velten et al. 2013; Wu et al. 2014b; Wu et al. 2014a], or a combination of source and camera (time-of-flight sensor) that are synchronously modulated at radio-frequencies in time [Heide et al. 2013; Kadambi et al. 2013; O'Toole et al. 2014b; Heide et al. 2014]. In the latter case, the performance critically depends on the method used to solve the so-called "multi-path interference" problem [Dorrington et al. 2011; Heide et al. 2014]. The equivalence of pathlength decomposition and light propagation has also been exploited for efficient rendering of light-in-flight [Jarabo et al. 2014].

Finally, there exist methods that provide simultaneous decompositions of space and pathlength. In the computer graphics literature, these fall into two categories. First are methods that decompose transport into a direct component ("one-bounce paths") and an indirect or global component ("multi-bounce paths") [Nayar et al. 2006; Gupta, M. and Agrawal, A. and Veeraraghavan, A. and Narasimhan, S.G. 2011; Reddy et al. 2012]. These imply a matrix $\mathbf{M}$ that has only one non-zero entry per row, and a spatially-adaptive function $w(\tau)$ that selects for each such non-zero entry the shortest pathlength τ for which transport is non-zero. In these cases, the pathlength resolution $\Delta\tau$ determines how clean the resulting separations are, with sharper resolutions allowing less of the global component—such as that caused by sub-surface scattering in Figure 1—to bleed into the direct one. Second is the temporal frequency probing method of O'Toole et al. [2014b]. This allows decompositions with more general choices of matrix $\mathbf{M}$ and function $w(\tau)$, for instance capturing transient images corresponding to fixed spatial probing patterns $\mathbf{M}$.

Interferometry. *Optical interferometry* refers to a large set of imaging techniques that exploit interference between one or more electromagnetic fields [Hariharan 2003]. Despite their popularity in other disciplines, interferometric techniques have rarely been used for computer graphics. An exception is Cossairt et al.'s [2014] use of a Michelson interferometer for refocusing.

Many interferometric techniques can be interpreted using Equation (3). Abramson [1983] used holographic techniques to create visualizations of light-in-flight similar to recent transient imaging; and the connection between interferometry and time-of-flight sensors has been discussed elsewhere [Schwarte et al. 1995; Luan 2001]. Here, we re-establish this relationship, and we expand it to relate interferometry to other types of light path decompositions.

Among interferometric techniques, the one most closely related is optical coherence tomography (OCT) [Huang et al. 1991], which is used for range scanning and visualizing volume cross-sections at micron resolutions. There are many variants based on single or multiple shots (Fourier-domain vs. time-domain OCT) and single- or full-field illumination. In our paper, we adapt a full-field, time-domain OCT configuration for broader computational imaging. We re-interpret conventional OCT measurements using Equation (3), and then build on this to measure other light transport decompositions with different $\mathbf{M}$ and $w(\tau)$ choices. We focus on time-domain OCT, as opposed to Fourier-domain, because it additionally allows decomposing transport according to wavelength and polarization.

Comparison and trade-offs. The main advantage of using interferometry to decompose transport is high pathlength resolution. Fempto-second lasers and time-of-flight sensors can achieve pathlength resolutions of about 600 μm and 10,000 μm, respectively. By comparison, we achieve pathlength resolutions of about 10 μm. Other interferometric techniques could be used to produce resolutions that are even higher, reaching sub-micron scales, for instance using very wide-band sources and post-capture processing that includes phase shifting. However, this would come at the expense of the ability to probe different spectral channels. Since we are interested in measuring wavelength-dependent effects, such as dispersion, birefringence, and scattering, we do not attempt to push pathlength resolution to those limits here.

This increased pathlength resolution is valuable in cases where it is necessary to obtain high-fidelity scans of small volumes, such as in laboratory material measurements or in scenes with natural microstructures. While macrophotography readily provides a way to image such scenes at *spatial* resolutions that approach the diffraction limit (a few microns), comparable pathlength resolutions cannot be achieved without the use of interferometry.

The main disadvantages of using interferometry to decompose transport are reductions in field of view, depth of field, and speed of capture. Our working volume can only accommodate material samples or small objects, and our system is not appropriate for table-top or room-scale scenes. As we explain in Section 3, our method operates by slicing a volume in micron-sized depth intervals. This means that measuring a 1 cm deep volume requires 10,000 images and several hours, during which the scene must remain still. This is exacerbated by the fact that our optical setup is very sensitive to vibrations induced by environment sources, and even micron-scale vibrations can severely affect our measurements.

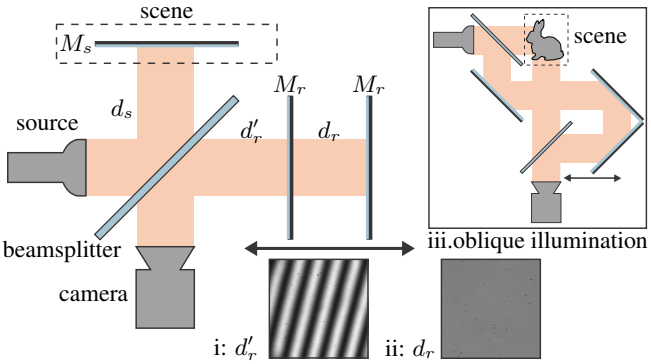


Figure 2: *Michelson interferometer. An input beam is split by a beamsplitter into two copies that traverse to mirrors M_r and M_s at distances d_r and d_s from the splitter, respectively. The two copies reflect back and recombine at the beamsplitter before being imaged. Insets (i) and (ii) show images for two positions of mirror M_r that either induce interference ($d'_r \approx d_s$) or do not ($d_r \neq d_s$). Inset (iii) shows a setup that can be used to achieve oblique illumination.*

Another limitation of our approach relative to non-interferometric techniques is reduced flexibility in the spatial probing patterns $\mathbf{M}$. We control these patterns by using mirrors of different shapes (Figure 6). This means that we can only induce patterns that correspond to manufacturable mirrors, and that changing the pattern $\mathbf{M}$ requires substantial manual reconfiguration. This is in stark contrast to spatial probing methods [O'Toole et al. 2012; O'Toole et al. 2014a; O'Toole et al. 2014b] that allow complex patterns and simple reconfiguration in software.

The third limitation of our approach is shared by all methods for pathlength decomposition, namely, the reduced signal-to-noise ratio when multiple paths with widely varying energies arrive at the same pixel. In extreme cases, this makes it difficult or impossible to resolve low-energy paths, and unlike conventional HDR photography, this cannot be solved by capturing multiple exposures.

3 Optics Background

We first present the background necessary for understanding interferometry and our interpretation of it as decomposing transport. Our description of temporal coherence follows Goodman [2000], and our description of spatial coherence follows Levin et al. [2013].

Interferometry begins with the Michelson interferometer, shown in Figure 2, with a light source of simultaneously small *temporal coherence* and *spatial coherence*. We will explain these notions and the operation of the Michelson interferometer intuitively, and then describe them analytically. The setup uses a beamsplitter to split the light wave emitted from the light source in two parts. One part is transmitted towards a reference mirror M_r placed on a translation stage. The other part is reflected towards the target scene, which for our intuitive explanation is another mirror M_s, but can generally be any arbitrary scene. We refer to the sides of the setup containing the reference and target mirror as the *reference arm* and *target arm*, respectively. Both waves are then reflected, recombined at the beamsplitter, and imaged by a camera. We denote by d_r, d_s the distance between the beamsplitter and mirrors M_r and M_s, respectively.

The small temporal coherence of the light source means that, when $|d_r - d_s|$ is larger than an amount called the *temporal coherence length* of the source, the two waves do not interfere and the camera will measure an image equal to the sum of the two images of the mirrors, as shown in Figure 2.ii. If instead we use the translation stage to reposition M_r to a position d'_r so that $|d_s - d'_r|$ is smaller than the coherence length, the two waves will interfere and the camera will measure a fringe pattern, as shown in Figure 2.i.

When we replace mirror M_s with a general scene, we can capture a sequence of images at different translations of the reference mirror

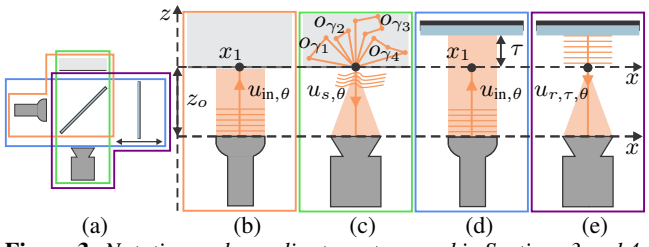


Figure 3: *Notation and coordinate system used in Sections 3 and 4. (a) The parts of the Michelson interferometer corresponding to the scene and reference arms. (b-e) The scene and reference arms shown unfolded in the same coordinate system. (b) and (c) show the target arms, with lighting and camera respectively. (d) and (e) show the reference arms, with lighting and camera respectively.*

M_r. By detecting interference patterns in the measured images, we can separate light paths of different pathlengths, up to a resolution equal to the temporal coherence length. We can additionally use sources with small *spatial* coherence to separate paths by location. This is because only light-paths that begin sufficiently close to each other on the source (i.e., are less than one *spatial coherence length* apart) and combine at the same sensor pixel will interfere. Note that this discussion is based on the coaxial interferometer design of Figure 2, which is the design we use exclusively in this paper. Oblique illumination is achievable, but it would require additional components as shown in Figure 2.iii.

This discussion highlights the importance of temporal and spatial coherence, which we will now describe analytically. They are wave phenomena, but scalar wave theory is sufficient to describe them, without having to employ the full complexity of electromagnetic theory. To simplify notation, and without loss of generality, we consider a two-dimensional world, with spatial points represented by their (x, z) coordinates. Figure 3 shows the coordinate system.

3.1 Temporal Coherence

We begin by assuming that the source produces a plane wave propagating in the z direction, so that the electromagnetic field is independent of x. Such a wave can be produced by transferring the output of a point source through a lens, as in Figure 8. If the wave is monochromatic at wavelength λ, the electromagnetic field at position (x, z) and time t can be described as

$$u(x, z, t) = a \cdot \exp\left(ik(-z\eta + ct)\right), \tag{4}$$

where c is the speed of light in vacuum, η is the refractive index of the medium, $k = 2\pi/\lambda$ is the wavenumber, and a is the complex amplitude of the source.

Real waves are not perfectly monochromatic, but are better described as superpositions of plane waves of different wavelengths,

$$u(x, z, t) = \int_k a_k \cdot \exp\left(ik(-z\eta + ct)\right)\, \mathrm{d}k, \tag{5}$$

where the complex number a_k is the amplitude of each component. Typical real world sources can be modeled [Goodman 2000] by assuming that a_k are random variables with phases sampled uniformly and powers $|a_k|^2$ sampled from a Gaussian with mean $\bar{k}$ and standard deviation Δ_k,

$$|a_k|^2 \propto e^{-\frac{(k-\bar{k})^2}{2\Delta_k^2}}. \tag{6}$$

The standard deviation Δ_k is the *spectral bandwidth* of the source: smaller values of Δ_k indicate a more monochromatic source.

Now suppose a sensor records at pixel x a measurement $I(x)$ of the superposition of two copies of the wave u from Equation (5) that have traveled different distances z and $z + \tau$. The measurement is

averaged over exposure time, producing

$$I(x) = \left\langle |u(x,z,t) + u(x,z+\tau,t)|^2 \right\rangle_t \quad (7)$$

$$= 2I_o(x) + 2\,\mathrm{Re}\left\{\mathrm{corr}(x,\tau)\right\}, \quad (8)$$

where $\langle\cdot\rangle_t$ denotes expectation over time, $I_o(x) = \left\langle |u(x,z,t)|^2 \right\rangle_t$ is the intensity of the original wave, and

$$\mathrm{corr}(x,\tau) = \langle u(x,z,t)^* \cdot u(x,z+\tau,t)\rangle_t \quad (9)$$

is the correlation of the two waves. If zero correlation exists between u and its shifted copy, the measurement will simply be equal to twice the intensity of the original wave, $2I_o(x)$. If positive correlation (constructive interference) or negative correlation (destructive interference) exists, the measurement will vary between 0 and $4I_o(x)$, depending on the relative shift τ of the two copies. The maximum value is achieved when there is perfect correlation ($\tau = 0$). We will refer to the real part of the field correlation, equal to the difference $I(x) - 2I_o(x)$, as the *interference contrast*. The *temporal coherence length* of the source, which we denote by L_c, is the largest value of τ for which significant correlation exists and therefore significant interference contrast is measured. To evaluate the temporal coherence length, we use the following claim.

Claim 1. *The field correlation magnitude is a Gaussian with standard deviation inversely proportional to the spectral bandwidth Δ_k,*

$$\mathrm{corr}(x,\tau) = e^{i\bar{k}\tau} G_{\Delta_k^{-1}}(\tau), \quad (10)$$

where

$$G_{\Delta_k^{-1}}(\tau) = e^{-\frac{1}{2}(\Delta_k \tau)^2}. \quad (11)$$

Before proving the above property, we consider and experimental visualization of the result. Figure 4 shows the interference component we measured using the Michelson interferometer setup of Figure 2, plotted as a function of pathlength difference. As shown in the close-up plot, the profile looks like a fringe pattern, which varies as a function of the phase difference introduced by τ. When $\bar{k}\tau = (2p+1)\pi$ (p any integer), we observe destructive interference (zero intensity); and when $\bar{k}\tau = (2p)\pi$, we observe constructive interference (twice the signal intensity). As shown in the zoomed-out plot, the contrast of the fringe pattern decays and eventually becomes almost zero once the pathlength difference τ exceeds some amount. We can define this amount as the temporal coherence length of the source, which from Claim 1 is the standard deviation of the correlation window, $L_c = 1/\Delta_k$. We observe that L_c is inversely proportional to spectral bandwidth. The larger the bandwidth of the source, the smaller the temporal coherence length, and therefore the higher the resolution at which we can discriminate between paths of different lengths.

Proof of Claim 1. To evaluate the correlation, we note from Equation (5) that the Fourier transform of $u(x,z,t)$ over the time domain is the signal $a_k e^{ikz}$. Similarly the Fourier transform of $u(x,z+\tau,t)$ is $a_k e^{ik\tau} e^{ikz}$. Using Parseval's theorem, the inner products in the temporal and frequency domains are equivalent,

$$\mathrm{corr}(x,\tau) = \left\langle \int |a_k|^2 \cdot e^{ik\tau}\,\mathrm{d}k \right\rangle. \quad (12)$$

Then, using also Equation (6), we have

$$\mathrm{corr}(x,\tau) = \int e^{-\frac{(k-\bar{k})^2}{2\Delta_k^2}} \cdot e^{ik\tau}\,\mathrm{d}k. \quad (13)$$

Equation (13) is (up to a constant) the Fourier transform of a Gaussian with s.t.d Δ_k. For a Gaussian centered at zero, the Fourier transform is a Gaussian with inverse s.t.d. $1/\Delta_k$. As the Gaussian is not centered at zero, this is multiplied by a sinusoid representing the phase shift, resulting in Equation (10). □

3.2 Spatial Coherence

So far we have considered the case of a perfect plane wave propagating along the z axis. Such a wave has perfect spatial correlation, in the sense that if we know the field value $u(x,z,t)$, we can also predict its value at any spatial shift $u(x+\xi,z,t)$. This is only realizable by using a lens to collimate the output of a source that has infinitesimal area. However, many real sources have larger effective areas. This induces another effect known as spatial incoherence [Levin et al. 2013], which also affects interference contrast.

Each point on an area source emits an independent wave, resulting cumulatively in a collection of independent plane waves $u_\theta(x,z,t)$ over a small angular range $\theta \in [-\Theta/2, \Theta/2]$. As in the previous section, we consider two copies of the cumulative wave, but in this case with one copy shifted not only in the z, but also in the x direction. The intensity at a pixel on the sensor is given by integrating the intensities from Equation (7) for all independent waves,

$$I(x) = \int_{-\Theta/2}^{\Theta/2} \left\langle |u_\theta(x,z,t) + u_\theta(x+\xi,z+\tau,t)|^2 \right\rangle_t \mathrm{d}\theta. \quad (14)$$

This implies that the correlation is,

$$\mathrm{corr}(x,\xi,\tau) = \int_{-\Theta/2}^{\Theta/2} \langle u_\theta(x,z,t)^* \cdot u_\theta(x+\xi,z+\tau,t)\rangle_t \,\mathrm{d}\theta, \quad (15)$$

and similarly for the field intensity $I_o(x)$ and the interference contrast $I(x) - 2I_o(x)$. To evaluate how fast the interference contrast decays as a function of spatial shift , we use the following claim.

Claim 2. *For an area light source with an angular range $\theta \in [-\Theta/2, \Theta/2]$, the correlation decays as*

$$\mathrm{corr}(x,\xi,\tau) \approx W_{\Delta_c}(\xi) e^{i\bar{k}(\tau)} G_{\Delta_k^{-1}}(\tau), \quad (16)$$

where

$$W_{\Delta_c}(\xi) = \mathrm{sinc}\left(\frac{\xi}{\Delta_c}\right),\ \Delta_c = \frac{\bar{\lambda}}{2\Theta}. \quad (17)$$

We see that, for area sources, the interference contrast decays very fast as the spatial shift ξ increases, and no significant contrast is measured for values of ξ larger than the quantity Δ_c of Equation (17). Analogously to the previous section, we define Δ_c as the *spatial coherence length*. The spatial coherence length is inversely proportional to the angular extent of the source. The wider the source, the smaller the spatial coherence length, and therefore the higher the resolution at which we can discriminate between paths with different points of origin on the light source.

Proof of Claim 2. A plane wave propagating in angle θ is

$$u_\theta(x,z,t) = \int_k a_{k,\theta} \cdot \exp\left(-ik(\sin(\theta)x + \cos(\theta)z - ct)\right)\,\mathrm{d}k, \quad (18)$$

where we assumed $\eta = 1$ for simplicity and without loss of generality. For small θ, $\sin(\theta) \approx \theta$ and $\cos(\theta) \approx 1$, therefore

$$u_\theta(x,z,t) = \int a_{k,\theta} \cdot \exp\left(-ik(\theta x + z - ct)\right)\,\mathrm{d}k. \quad (19)$$

The time delay between the shifted wave $u_\theta(x+\xi,z+\tau,t)$ and the original one $u_\theta(x,z,t)$ is the projection of the shift (ξ,τ) on the propagation direction $(\sin(\theta),\cos(\theta)) \approx (\theta,1)$,

$$u_\theta(x+\xi,z+\tau,t) = u_\theta(x,z,t-(\theta\xi+\tau)/c). \quad (20)$$

A small adaptation of Claim 1 shows that, for every θ,

$$\mathrm{corr}_\theta(x,\xi,\tau) = e^{i\bar{k}(\theta\xi+\tau)} e^{-\frac{1}{2}(\Delta_k(\theta\xi+\tau))^2}, \tag{21}$$

where the shift between the waves is now $\theta\xi + \tau$ instead of τ as in Claim 1. If this shift is large, the interference contrast for every θ decays as a Gaussian with width equal to the temporal coherence length $1/\Delta_k$. However, even if the temporal coherence length is large relative to the spatial shift, so that in Equation (21) $e^{-\frac{1}{2}(\Delta_k(\theta\xi+\tau))^2} \approx 1$, the observable interference still decays if ξ is larger than Δ_c. To see this, we compute the total interference contrast by integrating Equation (21) over angles in the source,

$$\mathrm{corr}(x,\xi,\tau) = \int_\theta e^{i\bar{k}(\theta\xi+\tau)} e^{-\frac{1}{2}(\Delta_k(\theta\xi+\tau))^2} \,\mathrm{d}\theta. \tag{22}$$

Though we can compute Equation (22), to obtain a simpler expression we approximate the Gaussian term by its central value,

$$\begin{aligned}\mathrm{corr}(x,\xi,\tau) &\approx e^{-\frac{1}{2}(\Delta_k\tau)^2} \int_\theta e^{i\bar{k}(\theta\xi+\tau)} \,\mathrm{d}\theta \\ &= W_{\Delta_c}(\xi) e^{i\bar{k}(\tau)} G_{\Delta_k^{-1}}(\tau).\end{aligned} \tag{23}$$

The last equality follows from the fact that the term $e^{i\bar{k}\theta\xi}$ is a sinusoid varying with θ, whose integral is a sinc. Therefore, for large ξ, integrating over θ can completely eliminate the correlation. □

3.3 Coherence Properties of Different Sources

There are many types of light sources with different coherence properties. As we will see in the next section (e.g. Figure 5), these properties determine the type of transport decomposition that our interferometer provides. We discuss various sources here, limiting our attention to those in the visible spectrum.

Laser sources are at one extreme end of coherence. These are high-power sources that can typically be modelled as ideal points (zero area), while also having very narrow, effectively monochromatic, spectral widths. As a result, their temporal and spatial coherence lengths can reach a few meters, implying poor spatial and pathlength resolution. For this reason, lasers are not appropriate for micron-scale light transport decompositions, but can be used when micron-scale resolution is not necessary [Abramson 1983].

At the other extreme are light-emitting diode (LED) sources. These can be either "colored", with a spectral bandwidth of a few tens of nanometers around their central wavelength, or very broadband, covering the entire visible spectrum. Necause of their lower power density, LED sources have larger emitting areas, with widths ranging from several millimeters to a few centimeters. As a result, they have micron-scale temporal and spatial coherence lengths. Their exact values can be controlled by modulating the source output with an optical color filter, and placing an aperture between the source and the collimating lens: A narrower spectral filter means longer temporal coherence, and a smaller aperture means larger spatial coherence length. The increase in coherence comes at the expense of lower light output, and therefore longer exposure time.

Between the two extremes, there are two source categories. The first category includes superluminescent diode (SLD) sources, and supercontinuum laser sources. Like *standard* lasers, these are high-power sources that have infinitesimal effective area, but at the same time, they are polychromatic like standard LEDs. SLDs have bandwidths comparable to colored LEDs, and supercontinuum lasers span the entire visible spectrum. Therefore, these sources combine the meter-size spatial coherence lengths of lasers with the micron-scale temporal coherence lengths of LEDs.

Finally, the opposite combination corresponds to a monochromatic area source. Such a source can be created using a spatially dense bundle of multiple pinhole laser sources, or using specially manufactured, large-area laser semiconductors (diode bars and stacks).

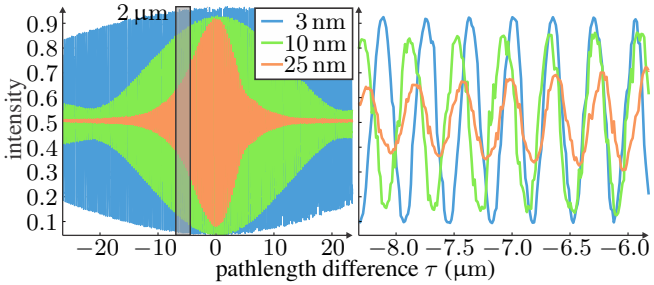


Figure 4: *Measuring the optical coherence length. The graph to the left shows the intensity measured by the camera in Figure 2 for different values of pathlength difference* $\tau = |d_s - d_r|$*, and the graph to the right shows a close-up over a range of* 2 µm*. When the pathlength difference is zero, maximal constructive interference exists. A pathlength shift of* $\lambda/2$ *results in destructive interference (zero intensity). As the pathlength difference increases, correlation between the waves is reduced and the fringe contrast decays, becoming almost zero once* τ *exceeds the temporal coherence length. The differently colored plots show measurements for different spectral bandwidths. Using a white light source with color filters of spectral bandwidth of* 3*,* 10 *and* 25 nm *produces temporal coherence lengths of* 50*,* 25*, and* 10 µm*, respectively. The measurements for this figure were taken at a resolution of* 10 nm.

Practical coherence values. To illustrate how these calculations translate into practice, consider the coherence lengths of an LED source coupled with a color filter and aperture.

With respect to temporal coherence length, note that broadband sources and color filters are often described by the width Δ_λ of the variation around the central wavelength $\bar{\lambda}$, rather than the variation Δ_k around the wave number. Using simple Taylor expansion, we can relate this to the temporal coherence length as

$$L_c = 1/\Delta_k \approx \frac{\bar{\lambda}^2}{2\pi\Delta_\lambda}. \tag{24}$$

Imagine a broadband LED, emitting in the entire visible range $[400, 700]$ nm. Ignoring any ultraviolet and infrared energy, we can set $\bar{\lambda} = 550\,\mathrm{nm}$ and $\Delta_\lambda = 150\,\mathrm{nm}$, resulting in coherence length $L_c \approx 0.3\,\mu\mathrm{m}$. In practice, this very fine resolution comes at the cost of very reduced interference contrast, due to chromatic aberration of optics. Alternatively, we can couple such a source with a color filter, or directly use a colored LED source. These alternatives give spectral widths in the range of $1 - 50\,\mathrm{nm}$, which correspond to temporal coherence lengths of a few microns. Additionally, as mentioned in Section 2, these narrower bands allow capturing measurements at multiple spectral channels. In Figure 4, we use the Michelson interferometer of Figure 2 to measure the temporal coherence lengths of such a combination, using three different color filters. Using a filter of bandwidth 25 nm, we measure a temporal coherence length of about $10\,\mu\mathrm{m}$. This is the resolution we use in most of our experiments in the paper.

In terms of spatial coherence length, as indicated by Equation (17), the values we can achieve in practice are a function of the angular extent of the source. For visible wavelengths, assuming that we can collimate the source output to have an angular range of $4°$, we have $\Delta_c \approx 8\,\mu\mathrm{m}$. With an angular extent of $1°$ the coherence length increases to $\Delta_c \approx 32\,\mu\mathrm{m}$. In a natural environment where illumination arrives from multiple directions over the hemisphere, coherence length can approach the wavelength of light.

4 Light Path Decomposition

Equipped with analytical representations of temporal and spatial coherence, we are ready to introduce a mathematical model of interferometric decomposition of light transport. The core of this section is Equation (34), which relates discrete interference measurements to an underlying pathlength-resolved transport function—a continuous version of the pathlength-resolved transport matrix of Equation (3). Using this model, we re-interpret the output of time-domain, full-field OCT as decomposition by pathlength τ (transient imaging), and we describe combinations of sources and reference mirrors that provide other types of decompositions.

Throughout this section, we assume that we image with a camera focused at depth z_o, as shown in Figure 3. As we restrict the discussion to points on the $z = z_o$ plane, from this point forward we simplify notation by omitting the z coordinate of fields. We denote by $u_{\text{in},\theta}(x,t)$ the incoming field arriving at both the reference and target arms at time t after being emitted from the source. This is a superposition of monochromatic waves, as in Equation (18). We denote by $u_{s,\theta}(x,t)$ the field scattered by the scene at space point (x, z_o) and time t. Similarly, we denote by $u_{r,\theta,\tau}(x,t)$ the reference field at point (x, z_o) when the reference mirror is positioned at depth $z_o + \tau/2$. We begin by deriving expressions that relate the reference and scattered fields to the input field.

Reference field. The reference field $u_{r,\theta,\tau}$ relates to $u_{\text{in},\theta}$ through a spatial transformation $f(x)$ and a temporal shift,

$$u_{r,\theta,\tau}(x,t) = u_{\text{in},\theta}(f(x), t - \tau/c). \tag{25}$$

In the standard Michelson interferometer of Figure 2, we have simply $f(x) = x$. We can produce more general transformations $f(x)$ with other mirror configurations (Figure 6). The form of $f(x)$ will determine the non-zero entries of the matrix $\mathbf{M}$ from Equation (3).

Scattered field. The scattered field $u_{s,\theta}$ relates to the input field through a more general transformation, which is a function of the scene's geometric, refractive, and material properties. To derive its form, we introduce some additional notation. We denote by $\{o_\gamma\}_{\gamma\in\mathfrak{A}(x_1,x_2)}$ the set of all possible paths in the scene from point (x_1, z_o) to point (x_2, z_o) (see Figure 3 for the case $x_1 = x_2$), by $\ell(o_\gamma)$ the *optical* length of each path, and by $\alpha(o_\gamma)$ the energy loss along the path—this loss is a function of the propagation distance and volumetric absorption along the path. Depending on the scene, such paths can be of many forms: direct paths traveling from the source to the object and then the camera, multiple reflections, multiple scattering, and so on. In the simple case that the scene is a perfect planar mirror, there are paths with non-zero absorption only if the start and end point are the same, $x_1 = x_2$, and for every x_1 only a single such path exists—the direct path from (x_1, z_o) to (x_1, z_o). For a path in air, where the refractive index $\eta \approx 1$, $\ell(o_\gamma)$ is equal to the geometric path length; in media with larger refractive indices, the path length also folds in the optical path delay. Finally, we denote by $\mathcal{T}(x_1, x_2, \tau)$ the complex scalar resulting from the integration of all paths from x_1 to x_2 with optical length τ,

$$\mathcal{T}(x_1,x_2,\tau) = \int_{\{o_\gamma|\ell(o_\gamma)=\tau\}_{\gamma\in\mathfrak{A}(x_1,x_2)}} \alpha(o_\gamma). \tag{26}$$

We call $\mathcal{T}(x_1, x_2, \tau)$ the *pathlength-resolved light transport function*. When the scene is a planar mirror, $\mathcal{T}(x_1, x_2, \tau) = \delta(x_1 - x_2, \tau - d_s)$, where d_s is the depth of the scene mirror. By definition, we observe that the function $\mathcal{T}(x_1, x_2, \tau)$ is the complex-valued continuous version of the pathlength-resolved light transport matrix, $\mathbf{T}^\tau$, introduced in Section 1. In parts of the following discussion, we will be employing the light transport matrix discretization for visualization and intuition, with the understanding that we continue to use the continuous version $\mathcal{T}(x_1, x_2, \tau)$.

We can now use the above functions to derive the scattered field $u_{s,\theta}(x, z_o)$. This is the superposition of contributions along all possible paths, with a time delay proportional to the path length,

$$\begin{aligned} u_{s,\theta}(x,t) =& \int_\xi \int_{\{o_\gamma\}_{\gamma\in\mathfrak{A}(x+\xi,x)}} \alpha(o_\gamma) u_{\text{in},\theta}\left(x+\xi, t - \frac{\ell(o_\gamma)}{c}\right) \mathrm{d}\xi \\ =& \int_\xi \int_\varsigma \mathcal{T}(x+\xi, x, \varsigma) u_{\text{in},\theta}\left(x+\xi, t - \frac{\varsigma}{c}\right) \mathrm{d}\varsigma \,\mathrm{d}\xi. \end{aligned} \tag{27}$$

The above equation relates the field to the function $\mathcal{T}(x_1, x_2, \tau)$ for the scene. The term $u_{\text{in},\theta}\left(x+\xi, t - \frac{\varsigma}{c}\right)$ is complex-valued and includes the phase difference between the integrated components, due to the different lengths of paths in each component.

Interference of reference and scattered fields. Having derived expressions for the reference and scattered fields, we now compute the interference contrast that can be measured by a camera. Ignoring for now diffraction blur, a camera focused at depth z_o will measure the superposition of the two fields,

$$\begin{aligned} I_{i,\tau}(x) &= \int_\theta \left\langle |u_{r,\theta,\tau}(x,t) + u_{s,\theta}(x,t)|^2 \right\rangle_t \mathrm{d}\theta \qquad (28)\\ &= I_s(x) + I_{r,\tau}(x) \\ &\quad + 2\,\mathrm{Re}\left\{ \int_\theta \left\langle u_{r,\theta,\tau}(x,t)^* \cdot u_{s,\theta}(x,t) \right\rangle_t \mathrm{d}\theta \right\}, \qquad (29) \end{aligned}$$

where by $I_s(x), I_{r,\tau}(x)$ we denote the intensities of the scattered and reference fields, respectively,

$$I_s(x) = \int_\theta \left\langle |u_{s,\theta}(x,t)|^2 \right\rangle_t \mathrm{d}\theta,\; I_{r,\tau}(x) = \int_\theta \left\langle |u_{r,\theta,\tau}(x,t)|^2 \right\rangle_t \mathrm{d}\theta. \tag{30}$$

The intensity components I_s and $I_{r,\tau}$ correspond to the images the camera would measure if we blocked the mirror or scene arms of the setup, respectively, and measured the other. By subtracting these components, we remain with the interference contrast, equal to the real part of the correlation of the reference and scattered fields,

$$C_\tau(x) = \int_\theta \left\langle u_{r,\theta,\tau}(x,t)^* \cdot u_{s,\theta}(x,t) \right\rangle_t \mathrm{d}\theta. \tag{31}$$

$C_\tau(x)$ is a correlation signal, analogous to $\mathrm{corr}(x, \xi, \tau)$ in Equation (15). We denote it by a distinct symbol because of its importance in our application. Using the coherence properties for correlation signals from Section 3, we can prove the following claim.

Claim 3. *The correlation of the reference and scattered fields is*

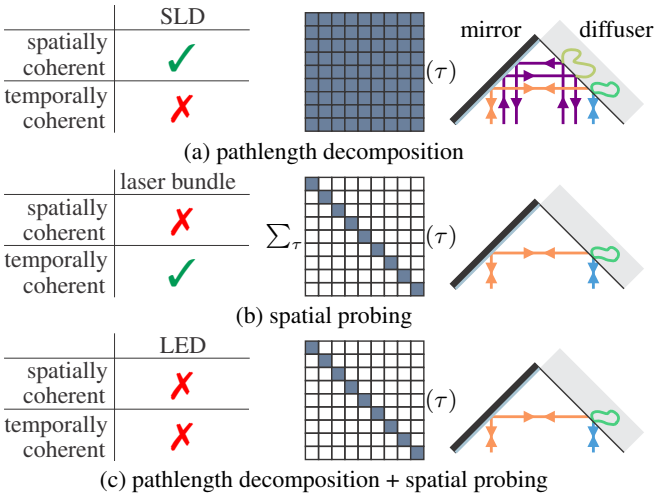
$$\begin{aligned} C_\tau(x) \propto& \int_\xi W_{\Delta_c}(\xi) \\ & \int_\varsigma \mathcal{T}(f(x)+\xi, x, \varsigma) e^{i\bar{k}(\varsigma-\tau)} G_{\Delta_k^{-1}}(\varsigma - \tau)\,\mathrm{d}\varsigma\,\mathrm{d}\xi. \end{aligned} \tag{32}$$

Proof. We denote $\xi = x + \zeta - f(x)$, implying $x + \zeta = f(x) + \xi$. Substituting Equations (25) and (27) in Equation (31), we obtain

$$\begin{aligned} C_\tau(x) =& \int_\xi \int_\zeta \mathcal{T}(f(x)+\xi, x, \varsigma) \\ & \int_\theta \left\langle u_{\text{in},\theta}\Big(f(x), t - \frac{\tau}{c}\Big)^* u_{\text{in},\theta}\Big(f(x)+\xi, t - \frac{\varsigma}{c}\Big) \right\rangle_t \mathrm{d}\theta\,\mathrm{d}\zeta\,\mathrm{d}\xi \end{aligned} \tag{33}$$

Using Claim 2, we get Equation (32). □

The above claim shows that the correlation $C_\tau(x)$ is equal to the pathlength-resolved light transport function $\mathcal{T}(f(x), x, \tau)$ from entrance point $(f(x), z_o)$ to exit point (x, z_o), blurred in space with a sinc of width proportional to the spatial coherence length, and blurred in the pathlength dimension with a Gaussian of width pro-



(a) pathlength decomposition

(b) spatial probing

(c) pathlength decomposition + spatial probing

Figure 5: *By equipping the Michelson interferometer with sources having different coherence properties, we can capture different light path decompositions. (a) Pathlength decomposition separates all paths in a scene ($\mathbf{M} = \mathbf{1}$), regardless of their endpoint locations, in terms of their optical length τ. It separates direct paths (blue), reflections (purple), retroreflections (orange), and all scattering (greens). (b) Spatial probing separates paths with certain start point locations while ignoring their lengths (by effectively measuring summations over pathlength). For a diagonal $\mathbf{M}$, it separates direct, retroreflection, and backscattering (dark green) paths. (c) Finally, it is also possible to combine the previous two cases and separate paths in terms of both pathlength and endpoint locations.*

portional to the temporal coherence length of the source. By recording the intensity $I_{i,\tau}(x)$ of Equation (28) and subtracting the individual intensities of the source and reference fields, we have

$$\begin{aligned} T^m(f(x), x, \tau) &\triangleq |I_{i,\tau}(x) - (I_s(x) + I_{r,\tau}(x))|^2 \\ &= |2\,\mathrm{Re}(C_\tau(x))|^2 \\ &= \left|2\,\mathrm{Re}\left(\mathcal{T}(f(x), x, \tau) * W_{\Delta_c} * (e^{i\bar{k}\tau} G_{\Delta_k^{-1}}(\tau))\right)\right|^2. \end{aligned} \tag{34}$$

This equation is our main result. It shows that using interference, we can measure blurred samples of the pathlength-resolved light transport function. The spatial resolution is controlled by the size of the spatial blur kernel, which is in turn controlled by the illumination angle of the light source. Similarly, the pathlength resolution is controlled by the size of the blur kernel in the pathlength dimension, which in turn is controlled by the spectral bandwidth of the light source. The exact correspondence $(f(x), x)$ between entrance and exit points where we sample the light transport function depends on the transformation implemented at the reference arm of the setup. Finally, the pathlength where we sample the light transport function depends on the depth of the reference arm.

4.1 Decomposition Types

Based on our model in Equation (34), the interferometer of Figure 2 can be equipped with various sources and mirror configurations to produce different types of light transport decompositions. The following discussion is summarized in Figures 5 and 6.

Pathlength decomposition. The first case is equivalent to transient imaging, and corresponds to setting $\mathbf{M} = \mathbf{1}$ and $w(\tau) = \delta(\tau - \tau_o)$ in Equation (3). For the former, we require a source with very large, essentially infinite, spatial coherence length, meaning that the spatial blur kernel W_{Δ_c} in Equation (34) is equal to 1 everywhere. For the latter, we need a source with a very short temporal coherence length, making the pathlength blur kernel $G_{\Delta_k^{-1}}$ in Equation (34) very narrow. As discussed in Section 3.3, this is achieved using an SLD or a supercontinuum laser. Capturing an entire transient sequence corresponds to sweeping over τ_o values, which can be done by densely scanning the reference arm to different positions. This hardware combination and capture procedure is exactly equivalent to conventional full-field, time-domain OCT.

Spatial probing. Referring again to Equation (3), another form of decomposition corresponds to setting, $w(\tau) = 1$ everywhere, and using a spatial probing pattern $\mathbf{M}$ with non-zero entries only for the paths we want to preserve. This requires a source with a near-infinite temporal coherence length and a very small spatial coherence length. From Section 3.3, such a source can be produced using a laser bundle. Additionally, we require a reference mirror configuration whose shape induces the spatial transformation $f(x)$ corresponding to the desired pattern $\mathbf{M}$. Using a planar mirror as in the standard Michelson interferometer corresponds to a matrix $\mathbf{M}$ that is non-zero only in its main diagonal. We discuss alternatives later in the section. Due to the infinite temporal coherence length, this case does not require a scan of the reference arm. We mention spatial probing here for completeness, but note that, for this type of decomposition, interferometry does not offer any advantages compared to the other, easier to implement and use, techniques discussed in Section 2. For this reason, we did not implement this case, and only used spatial probing combined with pathlength decomposition, as discussed in the next paragraph.

Pathlength decomposition and spatial probing. The last case decomposes light transport simultaneously in terms of pathlength and spatial correspondences. As suggested by the previous two cases, we can do this using a source with temporal and spatial coherence lengths that are both small. An LED source is the appropriate choice as described in Section 3.3. As is needed for pure pathlength decomposition, capture requires a dense scan of the reference arm to many different positions.

Creating probing patterns. In the last two decomposition types, replacing the planar mirror M_r in the reference arm of the Michelson interferometer with different mirror shapes induces different correspondences $f(x)$, that can be used to probe off-diagonal elements of the light transport matrix. For instance, the right-angle pair of mirrors in Figure 6(b) induces a vertical flip, or $f(x) = -x$. This corresponds to measuring the anti-diagonal of the light transport matrix, which includes two-bounce paths as shown in Figure 6(b). Shifting such a pair of mirrors vertically in the figure measures arbitrary anti-diagonals, as in Figure 6(c); and combinations of parallel mirrors and orthogonal mirror pairs sample different blocks of the light transport matrix in different ways, as shown in Figure 6(d). Note that the probing pattern $\mathbf{M}$ produced by these mirror configurations will be different if we change our setup so that the camera and source are no longer co-axial.

5 Pipeline and Design Considerations

Having presented our setup and methodology, we now discuss the computational post-processing and other practical considerations for optimizing the performance of our setup for graphics applications. Throughout this section, we consider a toy scene consisting of a tilted diffuse plane, as shown in Figure 7(a).

Capture and computational post-processing. Measuring the scene and mirror images I_s and $I_{r,\tau}$ separately is impractical, as it doubles acquisition time. In practice, since the reference arm contains a mirror configuration, the intensity $I_{r,\tau}$ does not vary much with τ and can be estimated by smoothing the τ dimension. For this reason, we perform a single scan of the scene, capturing a set of frames corresponding to $I_{i,\tau}(x)$ for various values of τ, as shown in Figure 7(a)-(b). Scanning a volume of 1 cm using this process typically requires capturing 10000 images (one image per micron-translation of the reference arm), a process which can take up to several hours for scenes requiring long exposure times. As we mention in Section 6, for scenes that require multiple spectral channels or high-dynamic range, we need to repeat the scanning process

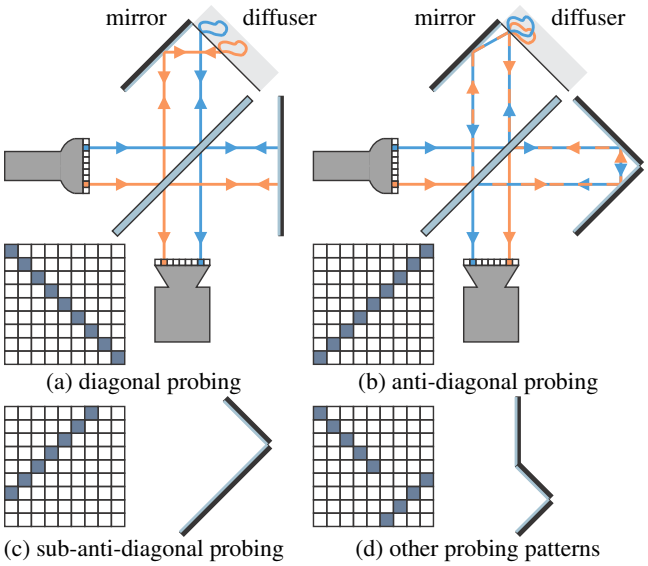


Figure 6: *Reference mirror configurations in a Michelson interferometer induce different spatial probing patterns. Assuming co-axial source and camera: (a) A planar mirror probes the main diagonal of the light transport matrix, corresponding to direct, retroreflection, and backscattering paths. (b) A right-angle mirror pair probes the anti-diagonal, which includes $2n$-bounce reflection paths. (c) Offsetting a right-angle pair allows probing away from the anti-diagonal. (d) Different combinations of these basic configurations probe blocks of the light transport matrix in different ways.*

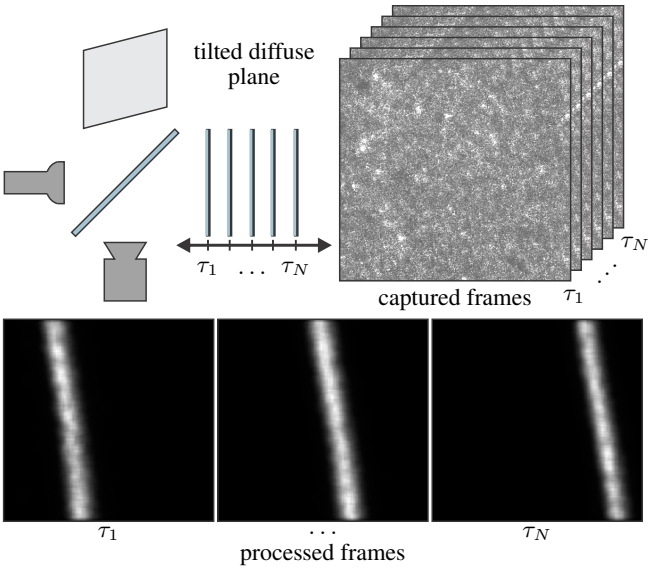


Figure 7: *Capture and computational pipeline. We consider a toy scene consisting of a diffuse plane, tilted relative to the optical axis of the camera and illumination, as shown in the upper left. By translating the reference mirror to different positions, we capture a sequence of frames corresponding to the intensity $I_{i,\tau}(x)$ measured for different values τ. By processing the frames using Equation* (36)*, we separate the interference component $\hat{T}^m(f(x), x, \tau)$, resulting in the images at the bottom. As the target plane is tilted, different points have different depths and therefore interfere with the target arm at different pathlength values τ, corresponding to the scan lines shown in the processed frames.*

several times, resulting in hundreds of thousands of images.

Following the capture section, in post-processing we approximate the interference-free component as

$$\hat{I}_s(x) + \hat{I}_r(x) \approx \text{median}_\tau(I_{i,\tau}(x)). \tag{35}$$

Furthermore, in real scenes the interference signal T^m of Equation 34 involves a high-variation pseudo-random speckle noise. To eliminate these speckle artifacts, we blur our measurements over a small spatial window with a filter $g(x)$. Our final estimate of the optical pathlength decomposition function is, then, computed as

$$\begin{aligned}\hat{T}^m(f(x), x, \tau) &= T^m(f(x), x, \tau) * g(x) \\ &= |I_{i,\tau}(x) - \hat{I}_s(x) - \hat{I}_r(x)|^2 * g(x).\end{aligned} \tag{36}$$

This blurring is one of the factors that result in a reduction in the spatial resolution of our measurements. We discuss speckle noise in more detail in the supplementary material. In the remaining of this section, we discuss other factors, as well as practices to maximize the signal-to-noise ratio of the measurements (36).

Contrast. Assume we want to capture the pathlength component which corresponds to optical length τ and that the reference mirror is positioned at the corresponding depth. We will define the contrast as the spatially averaged magnitude of the interference (correlation) component, over the intensity components,

$$\text{Contrast} = \frac{\sqrt{E_x\left[T^m(f(x), x, \tau)\right]}}{E_x[I_s(x) + I_{r,\tau}(x)]}. \tag{37}$$

To avoid saturation, we set the exposure time so that the averaged intensities without interference fall in the middle of the sensor's dynamic range, $E_x[I_s(x) + I_{r,\tau}(x)] \approx 0.5$. Then, Equation (37) is simply the averaged magnitude of the interference contrast.

A short calculation shows that the contrast in Equation (37) is maximized when the energy of the reference field $I_{r,\tau}(x)$ is equal to the energy of the scattered field at the optical length τ of interest $T^m(f(x), x, \tau)$. Because in diffuse scenes only a small portion of the energy in the target arm returns to the camera, to achieve the above ratio, we need to attenuate the very high amplitude of the reference arm. In most scenes, the intensity at a single point is the result of contributions from multiple paths, making it impossible to a-priori match the energy of the pathlength τ component only. Instead, we attenuate the reference arm so that its intensity matches the total intensity. As discussed in Section 6, we do this using either cross-polarization, or neutral density filters.

Optics blur. The ideal analysis of the Section 4 does not take into account blur introduced by the optics of the interferometry setup, including diffraction and sensor blur. These blur processes can reduce interference contrast, unless the various aperture and magnification parameters are chosen appropriately. We discuss these effects below, and in greater detail in the supplementary material.

We consider as before a camera focused at z_o, as shown in Figure 3. From diffraction theory [Goodman 1968], we have that instead of the field $u_\theta(x, t)$ arriving at point (x, z_o), the camera measures the field blurred by the camera diffraction blur kernel, $W_{\Delta_\Phi} * u_\theta(x, t)$, where the diffraction blur width $\Delta_\Phi = \bar{\lambda}/(2\Phi)$ is inversely proportional to the acceptance angle Φ of the camera.

The intensity measured by the camera is the intensity of the blurred field, averaged over the area of a pixel on the sensor,

$$I(x) = \Pi_{\Delta_x} * \int_{-\Theta/2}^{\Theta/2} \left\langle |W_{\Delta_\Phi} * u_\theta(x, t)|^2 \right\rangle_t \, \mathrm{d}\theta, \tag{38}$$

where Π_{Δ_x} is the pixel-sized rectangular function.

As a result of the above two processes, we prove in the supplemen-

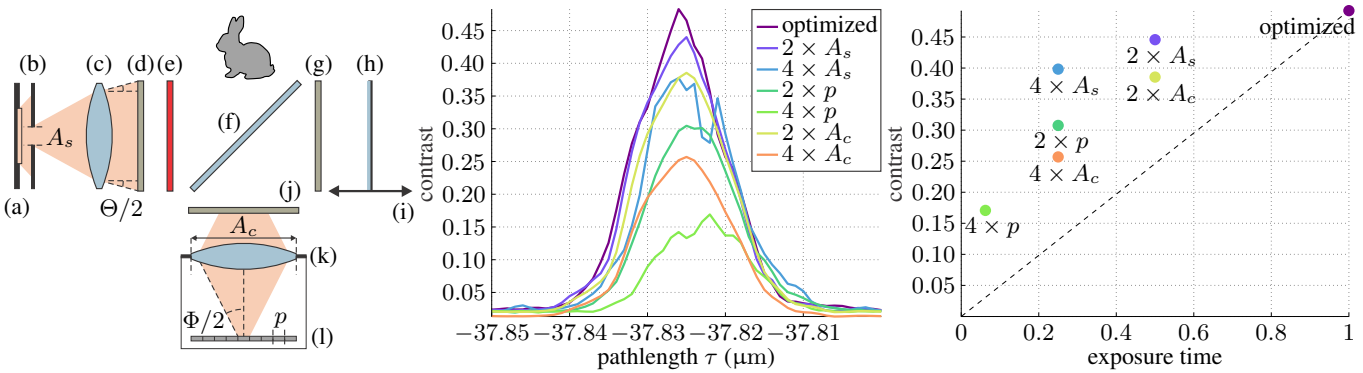


Figure 8: *Optimizing the Michelson interferometer of Figure 2. The schematic to the left describes our implementation. We consider the effect of the following imaging parameters: source aperture area A_s, camera aperture area A_c, and pixel pitch p. The optimized settings correspond to camera F-number of $f/32$, source aperture so that the source's angular extent is approximately half that of the camera, and pixel pitch of $4\,\mu\text{m}$. Different configurations are produced by modifying only one parameter. For each configuration, the middle graph shows the measured contrast for the tilted diffuse plane scene of Figure 7; and the right graph compares contrast and total exposure time.*

tary material that measurements from Equation (34) take the form,

$$T^m(f(x), x, \tau) = |2\,\mathrm{Re}\,(\beta * C_\tau(x))|^2 = \left|2\,\mathrm{Re}\left(\beta * W_{\Delta_c} * \mathcal{T}(f(x), x, \tau) * (e^{i\bar{k}\tau} G_{\Delta_k^{-1}}(\tau))\right)\right|^2, \tag{39}$$

where $\beta = \Pi_{\Delta_x} * W_{\Delta_\Phi}$.

It is important to observe that the additional blur β acts on the complex valued correlation component, rather than on powers (intensities). Therefore, if the correlation signal $C_\tau(x)$ contains spatial features whose frequency is higher than the widths Δ_x, Δ_Φ of the blur kernels, the averaged power of measurement signal is reduced. From Equation 32, we know that the signal $C_\tau(x)$ has limited spatial resolution Δ_c, as a result of the limited spatial coherence of the light source. Therefore, to achieve good contrast, we need to select the imaging parameters such that the optics blur remains below the signal resolution, that is, $\Delta_x < \Delta_c$ and $\Delta_\Phi < \Delta_c$.

We note that requesting $\Delta_\Phi < \Delta_c$ is equivalent to requesting that the angular extent of the source is smaller than the acceptance angle of the camera, $\Theta < \Phi$. In general, using a wide source is desirable, as it increases the overall power of the illumination and therefore reduces exposure time. However, increasing the source angular extent without adjusting the imaging parameters accordingly will also reduce the contrast of the correlation signal. To relate these directly to imaging parameters, we denote by Ψ the numerical aperture of the camera lens. A short calculation shows that a camera imaging at magnification m accepts light up to a maximal angle $\Phi = \frac{\Psi}{1-1/m}$.

In Figure 8, we quantify the above observations by measuring the detection contrast for the tilted diffuse plane scene of Figure 7 under different imaging configurations. We vary the size of the source and camera aperture, and the size of the camera pixel—which is analogous to changing the camera magnification. We observe that setting imaging parameters sub-optimally can result in loss of more than half the detection contrast. On the other hand, optimizing detection contrast comes at the cost of increased exposure time.

6 Implementation and Experiments

Figure 8 shows a schematic of our optical setup. We select sources and mirrors corresponding to configurations (a) and (c) of Figure 5, and (a) and (b) of Figure 6, respectively. The exact configuration depends on the application, as described below.

Our schematic has two additions compared to Figure 2. First, we modulate the output of the source source using color filters of bandwidth 25 nm (e), producing a temporal coherence length of about 10 μm (Figure 4). To enable RGB imaging, we use a filter wheel to automatically change between filters centered at different wavelengths: 625 nm (red), 525 nm (green), and 450 nm (blue).

Additionally, we use three polarizers: at the source (d), camera (j), and reference arm (g). The polarizers on the source and camera are rotated to be either parallel or crossed with each other, depending on which polarization component of the scene's transport is to be measured. The polarizer on the reference arm (g) is rotated relative to the other two in order to attenuate the reference beam by different amounts. This enables HDR imaging: we scan exposure brackets where, for each setting of the camera's exposure time, the reference polarizer (g) is adjusted so that the reference beam's intensity is at roughly one-quarter of the sensor's dynamic range (for a total intensity at roughly half the dynamic range when imaged together with properly exposed parts of the target scene). When we need to measure both polarization components, we remove polarizers (d) and (j), and replace (g) with a sequence of neutral density filters.

We discuss our implementation in detail in the supplementary material. In the rest of this section, we use this assembly to analyze transport in various scenes and for various applications. We first discuss experiments using an LED source (Figure 5(c)).

Depth scanning. Following Section 4.1, we use an LED source and a planar reference mirror to measure the pathlength decomposition of the diagonal component $T^m(x, x, \tau)$. As our camera and source are co-axial, we then obtain the depth $D(x)$ at pixel x as

$$D(x) = \min\{\tau : T^m(x, x, \tau) > 0\}. \tag{40}$$

Figure 9 shows scans of a few different objects. Each scan measures a volume of 15 mm at a step size of 1 μm, and requires roughly eight hours of capture time. The effective depth resolution is 10 μm. We can measure depth in a number of challenging situations, such as when materials exhibit substantial scattering (soap), semi-transparency (gummy bear), and strong caustics and specular or diffuse interreflections (cup and pasta). The very fine pathlength resolution of our setup translates directly into sharp depth resolution, which allows capturing details such as the hair texture on the coin and the fine texture on the gnocchi surface.

Figure 10 shows depth maps obtained for two more scenes. The stage step size and effective resolution are the same as before, but the volume and capture time increase to roughly 25 mm and 12 hours, respectively, to capture the longer reflection paths. The first scene includes sharp specular interreflections, with a single chess piece between an angled diffuse wall (right) and a mirror (left). In the acquired depth map, the mirror reflections are interpreted as real objects at a greater depth. Despite the indirect reflection by the mirror, the depth map of the chess piece is accurately recovered. The second scene shows a translucent gummy bear between two diffuse walls. Compared to the rightmost column of Figure 9, where

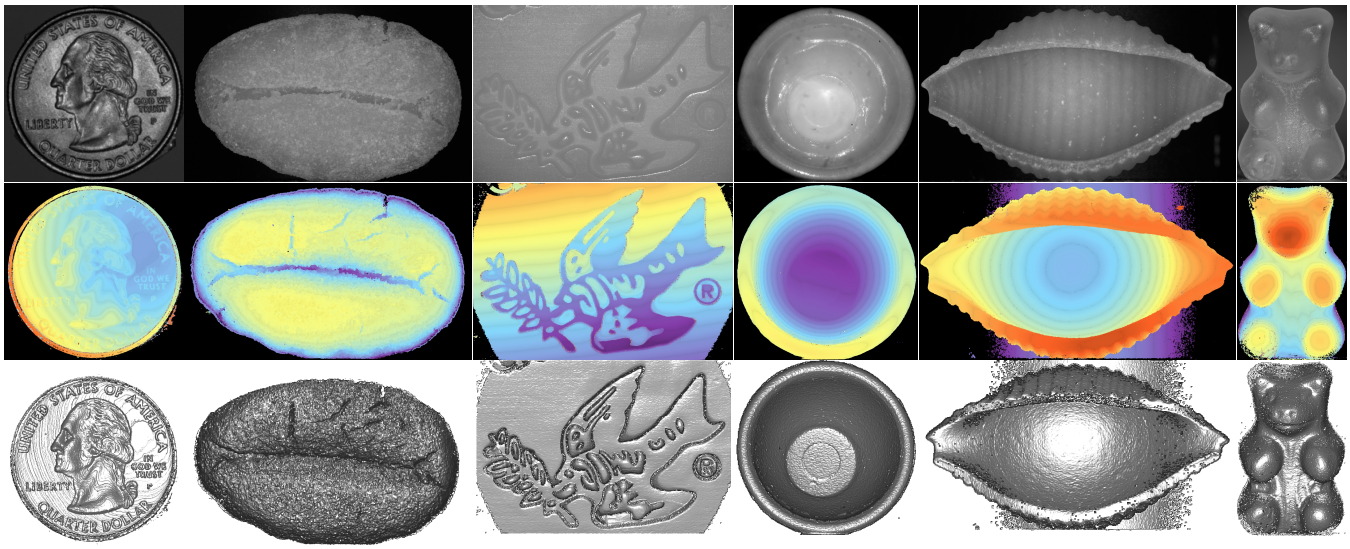


Figure 9: *Scanning depth. For each scene from top to bottom: an unprocessed HDR frame; scanned depth, visualized using pseudo-color and a periodic change in luminance for a contouring effect; and rendering of scanned 3D mesh. All scans are at a depth/pathlength resolution of* $10\ \mu\text{m}$*. Scenes from left to right: US quarter, gnocchi, soap carving, plastic toy cup, dry pasta, and gummy bear.*

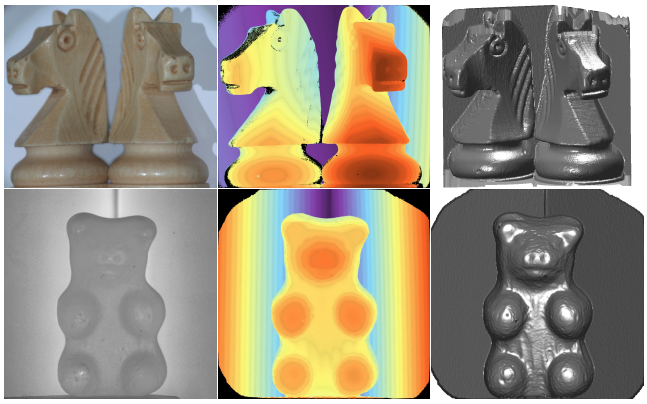


Figure 10: *Depth scans of complex scenes. From left to right: an unprocessed HDR frame; scanned depth, visualized as in Figure 9; and rendering of 3D mesh. Top row: Chess knight between a diffuse wall to the right and a mirror to the left. The mirror-reflection is interpreted as a real object at a greater depth. Bottom row: Gummy bear between two diffuse walls. The scene geometry is recovered accurately, despite the conflation of paths from the diffuse walls behind the near-transparent gummy bear.*

the gummy bear is surrounded by air, here the combination of a diffuse wall behind a near-transparent gummy bear induces many distinct paths of similar intensity at each camera pixel. Despite this conflation, the scene geometry is still recovered accurately.

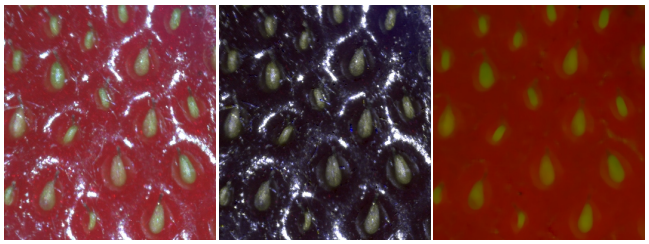
Separation of direct and global components. We can use the same source and mirror combination as before to also do direct-global separation. Since the source and camera are coaxial, the direct component at each pixel is the energy of the shortest path that contributes to the diagonal of the light transport matrix. Thus,

$$I_{\text{direct}}(x) = T^m(x, x, D(x)), \tag{41}$$

where $D(x)$ is the depth from Equation (40). The global component sums the contributions from all other paths and, given an image $I(x)$ captured using a conventional camera, can be obtained as

$$I_{\text{global}}(x) = I(x) - T^m(x, x, D(x)). \tag{42}$$

We note that spatial probing alone is not sufficient for computing



HDR frame direct component global component

Figure 11: *Direct-global separation for a close-up of a strawberry.*

the exact direct component, and it is necessary to also do pathlength decomposition. With reference to Figure 1, spatial probing (diagonal of the light transport matrix) captures the various blue paths, which include direct paths and retroreflections, but also backscattering. Using pathlength decomposition to select only the shortest path additionally removes the backscattering contributions.

Figure 11 shows such a decomposition for a close-up of a strawberry. We produced this separation from 12 scans of the strawberry (4 exposures $\times$ 3 spectral channels). Each scan was performed at a step size of $10\ \mu\text{m}$, meaning that we slightly undersampled in the pathlength domain relative to the source's pathlength resolution of $10\ \mu\text{m}$. Additionally, we used 8×8 pixel binning on the camera to speed-up frame streaming, and large source and camera apertures to decrease exposure time. Even though these imaging settings are suboptimal (see discussion of Figure 8), they were necessary to reduce the total scanning time to a little below two hours, thereby accommodating the limited shelf life of the strawberry.

We observe that the direct component is essentially achromatic and captures features characteristic of direct illumination, such as the specular highlights and high-frequency surface structure. The indirect component in this image is primarily volumetric scattering, and we see that it accounts for the red color of the strawberry and a large part of the coloration of the strawberry seeds.

Scattering. As discussed in Section 4.1, we can also obtain a decomposition of scattering paths that begin and end at the same locations on the surface of a scattering volume. To demonstrate this, Figure 12 shows the results from scanning a scene with materials of different scattering properties, such as a metal, a diffused wall and a highly-scattering zirconia coating. Figure 12(c) shows a processed frame from the scan, where the intensity tails at the zirconia coat-

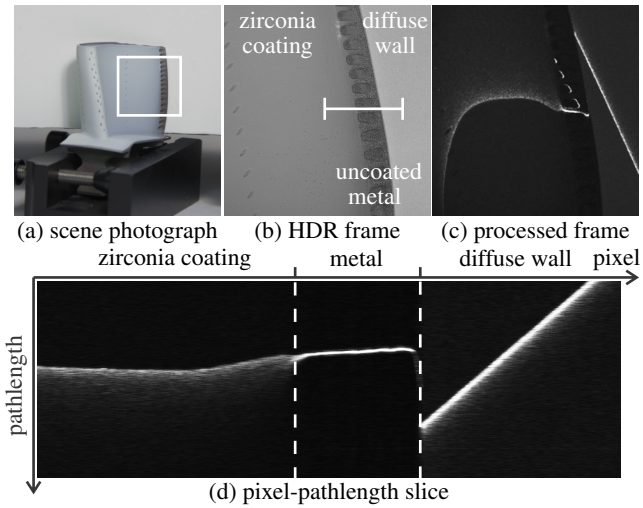


Figure 12: *Scanning a scene with materials of different scattering properties using an LED source. (a) A photograph of the scene: A turbine blade with a highly diffusive zirconia coating is placed in front of a spectralon slab. (b) An unprocessed frame. (c) A processed frame, showing as an image the value of the measurement $T^m((x,y),(x,y),\tau_o)$ for some value of τ_o and all pixels (x,y). (d) The function $T^m((x,y_o),(x,y_o),\tau)$ for pixels (x,y_o) on the horizontal crosssection shown in (b), and for all values τ.*

ing and diffuse wall regions are produced by the strong backscattering of those materials. In Figure 12(d), we see the pathlength slice $T^m((x,y_o),(x,y_o),\tau)$. (We use pairs (x,y) to indicate the 2D pixel coordinates). At pixels corresponding to zirconia coating and the diffuse wall, the interference remains noticeable over a long range of pathlengths. Conversely, we see negligible scattering in the uncoated metal pixels.

In Figure 13(a), we show scattering profiles $T^m((x_o,y_o),(x_o,y_o),\tau)$, as a function of pathlength. These were measured by scanning uniform slabs of various translucent materials: whole milk, milk soap, wax, and the zirconia coating mentioned previously. For materials that are optically dense and strongly forward scattering, such as whole milk, the measured intensity drops very sharply. On the other hand, for the diffuse zirconia coating, we can measure non-zero interference for pathlengths as long as 3 mm inside the medium.

In Figure 13(b), we additionally show the scattering profiles measured in the zirconia material for different polarization channels. Specifically, we measure the scattering profiles when the polarizers (d) and (j) in Figure 8 are rotated so that their axes are parallel or orthogonal ("crossed"). We observe at short pathlengths, the component corresponding to parallel polarization is much stronger than that corresponding to crossed polarization. However, as the pathlength increases, the parallel- and cross-polarized components become equal. This is in agreement with the well-known observation that low-order scattering preserves polarization, but high-order scattering becomes randomly polarized. For reference, we also plot the scattering profile for this material from Figure 13(a), where there is no decomposition of polarization.

Dispersion and birefringence. As discussed in Section 4, interferometry produces light path decompositions in terms of *optical* instead of geometric pathlength. Two light paths that have the same geometric length can have different optical length if they travel through media with different refractive indices. This could be either because they travel through different optical materials, or because of phenomena such as dispersion and birefringence, where the refractive index of a material changes as a function of the wavelength and polarization of light. We visualize these three phenomena using our setup with an LED source and a planar reference mirror.

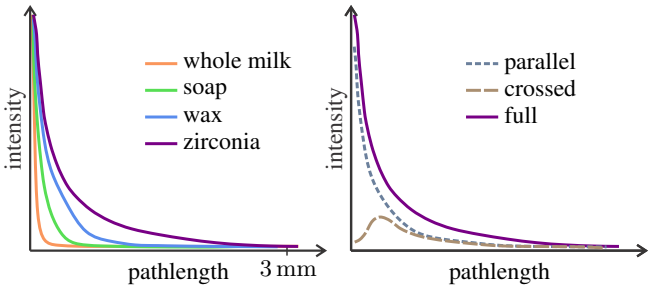


Figure 13: *Pathlength decomposition of scattering paths starting and ending at a single point on the surface of a scattering volume. All measurements use an LED source. Left: Measured scattering profiles $T^m((x_o,y_o),(x_o,y_o),\tau)$ at a single pixel (x_o,y_o) as a function of τ, for various materials. Right: Measured profiles for the parallel and cross-polarized scattering component of zirconia.*

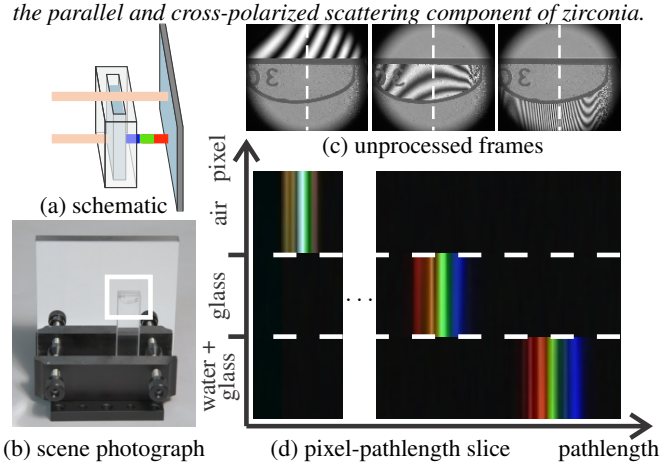


Figure 14: *Visualizing dispersion in the pathlength domain. (a) A schematic of the scene, consisting of a glass container partially filled with water and placed in front of a mirror. As light travels through the glass and water, different wavelengths are delayed by different amounts due to dispersion. (b) A photograph of the scene we scan with an LED source. We consider the close-up shown in the box. (c) Frames acquired at a wavelength of 525 nm, and at three different pathlengths. Spatial fringes indicates interference of paths that travel through, from left to right, air only, air and glass, and air, glass, and water. (d) Visualization of the interference measured across a vertical cross-section of the frames, as a function of optical pathlength and for three different wavelengths. We observe that paths travelling through glass or through both glass and water are delayed more compared to paths travelling through air only. Additionally, the delay is a function of wavelength, resulting in the appearance of a "rainbow" as a function of optical pathlength.*

In Figure 14, we scan a scene consisting of a glass container partially filled with water, placed in front of a mirror. Paths from the light source to the mirror and back to the camera all have the same geometric length. However, they have different optical lengths, depending on whether they travel through only air, through the empty part of the glass container, or through the part filled with water. Additionally, optical lengths vary as a function of wavelength, because of the dispersion of glass and water. This produces a "rainbow" when we visualize the measurements $T^m((x_o,y),(x_o,y),\tau)$ as a function of pathlength, even though such a rainbow is not visible in a static image. The path delay between different spectral bands increases for paths going through both water and glass (roughly 20 µm), compared to paths going through glass only (roughly 10 µm), due to the additional dispersion of water.

In Figure 15, we visualize birefringence by scanning a scene consisting of a plastic protractor placed in front of a mirror. The mechanical stresses in different parts of the plastic material result in

the refractive index of the material varying spatially as a function of polarization and wavelength, a behavior known as photoelasticity. This creates familiar color patterns when the protractor is viewed under standard polarized light without any interference (Figure 15(b)). The refractive index variations additionally mean that paths travelling through the protractor to the mirror and back to the camera will have different spatially varying pathlength. This is shown in Figure 15(d), where we scanned and visualized the contributions from different pathlengths at different parts of the protractor. The spatial variations of the measured optical pathlengths are smaller than $50\ \mu\text{m}$, and they are roughly aligned with the RGB color patterns that appear under standard polarized light.

Using different probing patterns. As discussed in Section 4.1, by using different reference mirror configurations, we can probe different components of the light transport matrix. In Figure 16, we compare measurements of the same scene with two of these configurations. The scene is a realization of the two-dimensional mirror-diffuser 90° corner we have been using as example throughout Section 4.1 (Figures 5 and 6), with the addition of a thin slab of glass bisecting the angle between them. We do all measurements using an LED source (Figure 5(c)) with our three spectral filters.

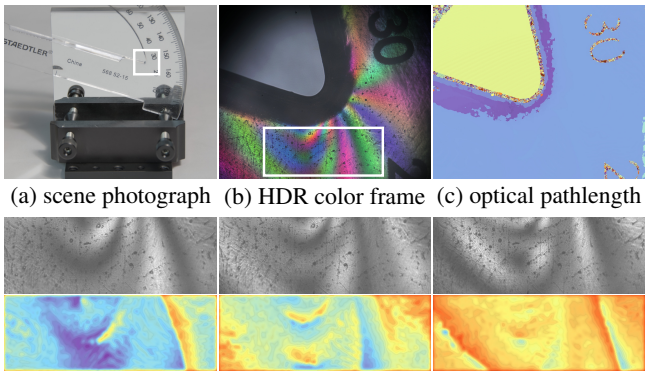
We first use a planar mirror in the reference arm, which allows measuring the pathlength decomposition of the primary diagonal of the scene's light transport matrix (Figure 6(a)). The left part of Figure 16(b) shows the measurements for pixels on a horizontal cross-section through the scene, as a function of pathlength. The slanted white stripe in the diffuser region of these measurements corresponds to direct paths (blue in the schematic of Figure 16(a)), whose pathlengths increase steadily toward deeper parts of the diffuse wall. Similarly, the slanted rainbow stipe in the mirror region corresponds to longer, retroreflection paths (orange in the schematic) that are addtionally affected by the spectral dispersion of the glass slab. The right of Figure 16(b) shows 2D image frames from the corresponding transient sequence at two different pathlengths. When played as a transient video, one sees the white stipe move right to left along the diffuser wall, and then sometime later, the rainbow stripe move right to left along the mirror wall.

Next, we place a right-angle mirror pair on the reference arm to measure the anti-diagonal component (Figure 6(b)). This component includes paths whose starting and ending locations are symmetric around the glass slab, such as mirror-glass-diffuser and diffuser-glass-mirror paths (purple in the schematic of Figure 16(a)). All such paths in this configuration have the same length, producing a horizontal line in the pixel-pathlength visualization of Figure 16(c). This implies that all frames of the corresponding transient sequence are zero, except for a single frame at the common length τ_o of the reflection paths, which is non-zero at all pixels. This frame is shown in the right part of Figure 16(c).

The spread of the rainbow patterns that are induced by dispersion in the glass slab are approximately $15\ \mu\text{m}$ for reflection and $30\ \mu\text{m}$ for retroreflection paths. The frame in Figure 16(c) corresponds to the green horizontal line in x-τ space. When played as a video, the transient sequence has analogous frames for the other two colors.

Comparing different pathlength decompositions. We now consider using an SLD source (Figure 5(c)). As discussed in Section 4, allows measuring pathlength decompositions of the entire light transport matrix, instead of only specific components of the matrix as in the LED case. In Figure 17, we use the plastic toy cup scene from Figure 9 to compare the two cases.

Figure 17(c) shows a pixel-pathlength slice for the dotted vertical line of pixels through the visual field shown in (a). This was measured using an LED source (Figure 5(c)) and corresponds to a pathlength decomposition of the light transport diagonal. It includes only direct reflection and backscattering paths—paths that, after scattering, exit the object from the same point at which they entered. Because interreflection and most scattering paths are removed, this decomposition facilitates measuring the 3D shape of the object through Equation 40, as shown in Figure 9.



(a) scene photograph (b) HDR color frame (c) optical pathlength

(d) Frame (top) and optical pathlength (bottom) detail (R, G, B)

Figure 15: *Visualizing photoelasticity. (a) We use an LED source to scan a scene consisting of a plastic protractor in front of a mirror. We consider the close-up shown in the box. (b) A standard RGB frame under polarized light with no interference or scanning. The color patterns on the protractor correspond to wavelength and polarization-dependent variations of refractive index due to photoelasticity. (c) Low-resolution ($1\ \text{mm}$) measurement of optical pathlength at every pixel, shown color-coded. (d) Zooming on the protractor, top row shows (left to right) the red, green and blue channels from (b); bottom row shows optical pathlength measurements for the same three channels and at a $10\ \mu\text{m}$ pathlength resolution.*

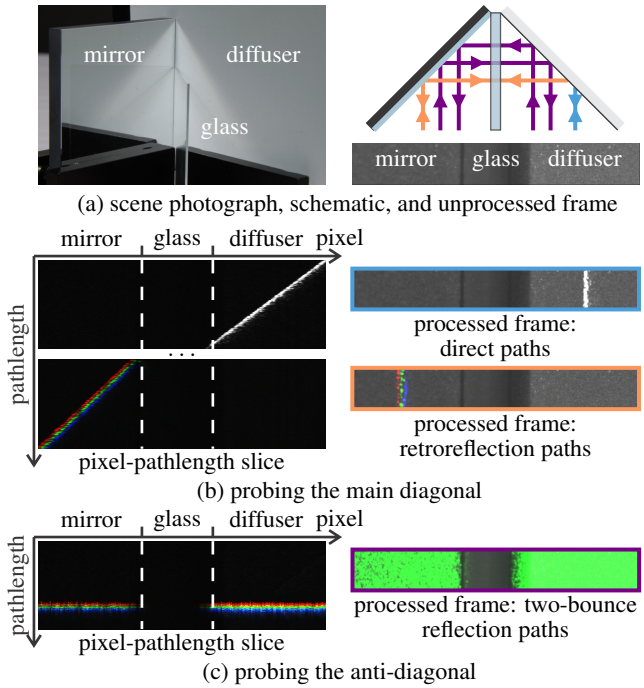
Figure 17(b) shows the analogous pixel-pathlength slice measured using an SLD source (Figure 5(a)). These measurements correspond to a pathlength decomposition of the entire light transport for this scene. Comparing 2D image frames from the two decompositions at the same pathlength (orange insets), we observe that in the SLD case, a much larger part of the frame has high intensity. This is due to the large number of additional scattering paths, other than backscattering, included in the measurements. Moreover, non-zero measurements are present at much longer pathlengths, corresponding to subsequent caustics and specular interreflections.

Light-in-flight visualizations. As discussed in Section 4.1, the pathlength decompositions measured by our setup combined with an SLD source, when composed into a video, are equivalent to light-in-flight visualizations produced by conventional transient imaging. Figure 18 shows representative frames from such visualizations for two miniature scenes. An additional example is shown in Figure 17(a), for the toy cup scene. The field of view in each scene is roughly $2\times2\ \text{cm}^2$, and the sequences are captured at a pathlength resolution of $10\ \mu\text{m}$, though for visualization they have been downsampled to $20\ \mu\text{m}$. This corresponds to a temporal resolution of $70\ \text{fs}$ in vacuum, or 15 trillion frames per second. The project website includes full sequences for these and additional scenes.

The scenes showcase different transport phenomena. The chess knight scene shows qualitatively different paths, including three types of reflections. In the zirconia scene, the central zirconia coating remains bright for several frames, due to the high amount of subsurface scattering. There is also transmission from the rear diffuse walls, through the ground glass that surrounds the central zirconia coating, and back toward the camera. Finally, at the vertex of the diffuse walls, we can more intense diffuse interreflections. The high pathlength resolution of our setup allows us to capture light-in-flight sequences with very crisp and well-localized wavefronts.

7 Conclusion

We have presented a computational imaging system that decomposes light transport based on spatial layout, pathlength, wavelength, and polarization. The system uses interferometry, and produces decompositions at a pathlength resolution of $10\ \mu\text{m}$. We have



(a) scene photograph, schematic, and unprocessed frame

(b) probing the main diagonal

(c) probing the anti-diagonal

Figure 16: *Probing different light transport matrix components. (a) We use an LED source to scan at three wavelengths a scene consisting of a thin glass slab placed between a mirror and a diffuser. A frame is shown at the bottom-right. The scene includes three types of paths: direct paths (blue), mirror-diffuser-mirror retroreflection paths (orange), and two-bounce reflection paths (purple). (b) Measurements of the primary diagonal of the light transport matrix, corresponding to direct and retroreflection paths. The left visualizes the measurements $T^m((x, y_o), (x, y_o), \tau)$ over pathlength τ for a horizontal line of pixels (a pixel-pathlength slice). The right shows image frames $T^m((x, y), (x, y), \tau_o)$ for all pixels (x, y) at two pathlength values. (c) Measurements of the anti-diagonal of the light transport matrix, corresponding to paths with start and end locations symmetric around the glass slab. These are two-bounce reflection paths that, for each wavelength, all have the same optical length. This produces horizontal stripes in the pixel-pathlength slice $T^m((x, y_o), (-x, y_o), \tau)$ (left), and a short sequence of all-bright frames $T^m((x, y), (-x, y), \tau_o)$ in the transient video (right, shown for the green wavelength). The rainbow patterns in (b) and (c) correspond to optical pathlength differences of* $15 - 30\,\mu\text{m}$ *between different wavelengths, due to dispersion in the glass slab.*

analyzed the various trade-offs in optimizing the system's performance, and used it to scan scenes with various optical phenomena.

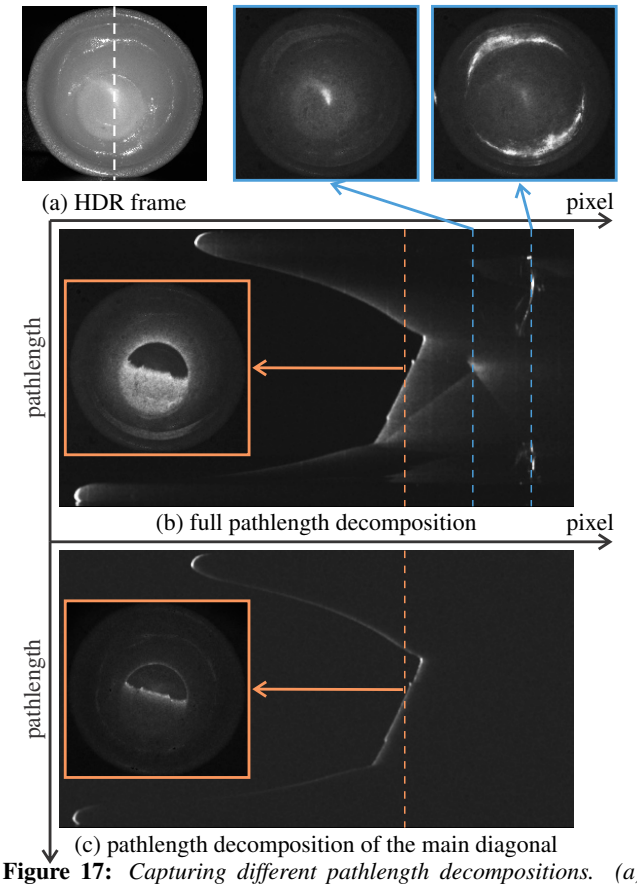
The high-fidelity decompositions produced by the system may be useful for a variety of direct and inverse problems in computer graphics and computer vision. In particular, we are intrigued by the challenges of developing inverse rendering algorithms for shape and material capture that take advantage of the information available in these high-resolution light path decompositions.

Acknowledgments

We thank Prof. David Clarke for providing the zirconia samples. Funding was provided by the NSF (IIS 1161564, 1012454, 1212928), ERC, ISF, Intel ICRI-CI, and AWS research grants.

A Speckle Analysis

The speckle properties of the images resulting from the interferometry setup described above are different than those obtained with standard imaging systems. Since speckles significantly effect the



(c) pathlength decomposition of the main diagonal

Figure 17: *Capturing different pathlength decompositions. (a) We scan a scene consisting of a plastic toy cup, and show pixel-pathlength slices for a vertical cross-section of pixels through the visual field. (b) When we capture a pathlength decomposition of the entire light transport matrix for this scene using an SLD, the pixel-pathlength slice includes direct paths, scattering, caustics, and specular interrreflections. These are also apparent in the 2D image frames of the transient sequence (orange and blue insets). (c) When we measure the pathlength decomposition only of the primary diagonal of the light transport matrix using an LED, the pixel-pathlength slice has only direct and backscattering paths.*

contrast of the interference signal, it is important to understand them and for this reason we analyze speckle properties below. We begin by discussing speckles under point light source illumination, then in Section A.2 we discuss area light sources where spatial incoherence needs to be taken into account. For clarity, we also review the source of speckles in standard imaging. To simplify notation, throughout this section we assume that we are using the standard Michelson interferometer setup of Figure 2, and therefore we do not need to use the spatial transformation $f(x)$. **Random phase statistics** To better explain the foundation of speckles, we consider the following toy problem. Let $\{\psi_i\}_{i=1}^n$ be n i.i.d random variables drawn from a uniform distribution on the $[0, 1]$ segment, and consider the random variable

$$\Psi = \left| \sum_{i=1}^{n} \frac{1}{n} e^{2\pi i \psi_i} \right|^2, \tag{43}$$

which is the mean of n variables with unit magnitude and independent phase. Since $E[e^{2\pi i \psi_i}] = 0$ and $|e^{2\pi i \psi_i}| = 1$, as the mean of n i.i.d. variable it is easy to show that $E[\Psi] = \frac{1}{n}$. Some additional calculation shows that the variance $V[\Psi] \approx \frac{1}{n^2}$.

That is the standard deviation of the random variable Ψ has the

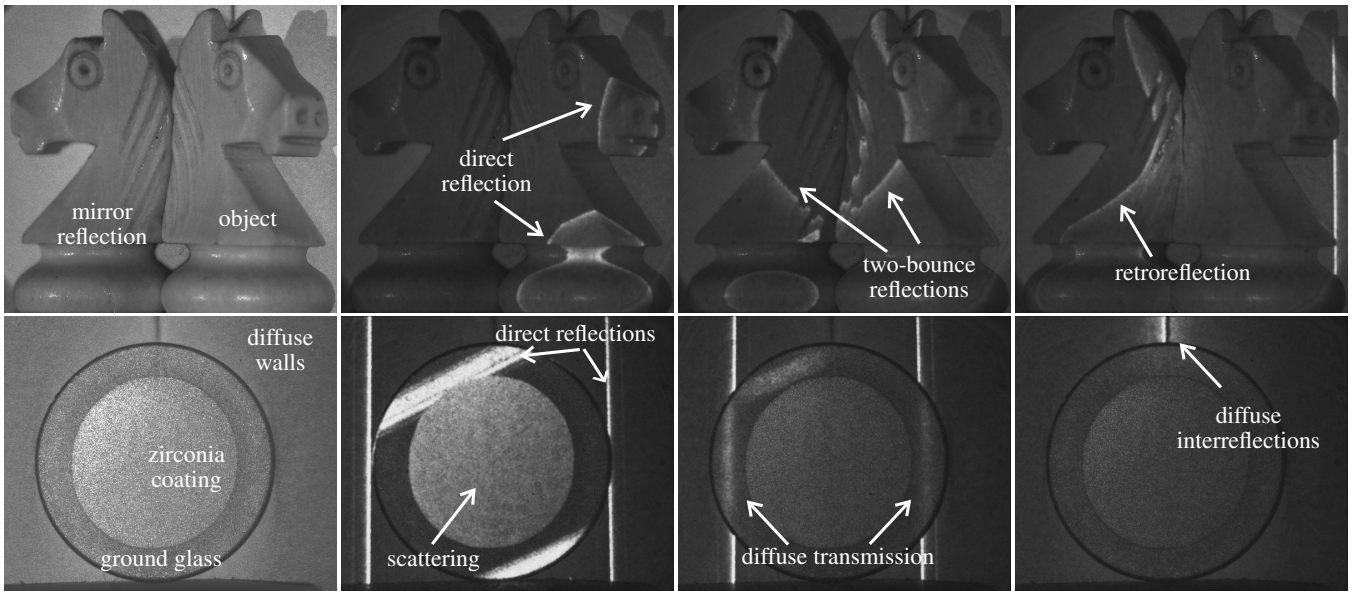


Figure 18: *Light-in-flight visualizations for two different scenes using an SLD source. For each row, the leftmost column shows a conventional photo of the scene, and columns two to four show representative frames. From top to bottom, the scenes are: chess knight between diffuse wall and mirror; zirconia layer on ground glass plate between diffuse walls.*

same magnitude as the signal itself. In other words, when we average instances with unit magnitude and random phase, we can compute the expected magnitude, but the magnitude of every particular instance we draw can deviate a lot from the expected one. As we show below this property is the hart of the speckles phenomenon.

A.1 Speckles Under Point Light Sources

A.1.1 Speckles in Standard Imaging

Consider the setup of Figure 3 with a camera focused at plane $z = z_0$. We denote by m the camera magnification between world coordinates to sensor coordinates (using ideal geometric optics ray mapping, world coordinate (x, z_o) is mapped to sensor coordinate $x' = mx$, m is usually negative). We also denote by Φ the maximal acceptance angle at this magnification.

Diffraction theory [Goodman 1968] implies that the intensity reaching a camera sensor is the convolution of the field with the diffraction blur kernel, averaged over the exposure length:

$$I^o(x) = \left\langle |W_{\Delta_\Phi} * u(x,t)|^2 \right\rangle_t \tag{44}$$

The diffraction blur kernel is a sinc whose width Δ_Φ is proportional to the numerical aperture and the magnification. Note that the exact width of the diffraction blur is wavelength dependent, however, to simplify the analysis we assume here that the spectral width of the sources we use is relatively narrow and neglect this dependency.

In Equation (44) the superscript ^{o} denotes a signal before sensor resolution is taken into account. In practice the sensor also has a finite pixel size. Let us denote by Δ_x the surface area being imaged into a single pixel. The sensor measures the infinite precision intensity I^o averaged over time and over the pixel area:

$$I(x) = \Pi_{\Delta_x} * I^o(x) \tag{45}$$

where Π_{Δ_x} denotes a rect filter of width Δ_x. Note that there is a significant different between the two blur processes, while the diffraction blur W_{Δ_Φ} operates on the complex field, the pixel blur Π_{Δ_x} operates on the intensities (powers).

For rough surfaces (which usually represent diffused materials) the local phase of the field $u(x,t)$ changes rapidly implying that $I^o(x)$ is a random variable with properties similar to the random variable Ψ define in Equation (43): it is the magnitude of the average of random phase elements. Therefore even if we can compute the expected intensity $E[I^o(x)]$, the diffraction limited intensity $I^o(x)$ we will measure at different spatial points has a very high variation, resembling a pseudo random pattern. As discussed above, the standard deviation can be as high as the mean intensity of the signal. This random intensity property is the source of the noisy speckles profile that one often observes. Note that the speckles width is controlled by the width of the diffraction blur W, which scales as a function of the numerical aperture- a large numerical aperture leads to narrow speckles and a small numerical aperture results in wide speckles. In practice the pixel area further blurs the actual measured intensity. Since the blur now operates on the intensity rather than the complex field, the speckles contrast is reduced. If the pixel area is much wider than the diffraction blur width $\Delta_x \gg \Delta_\Phi$ then by Equation (45), $I(x)$ averages multiple independent speckles $I^o(x)$. By the law of large numbers, this averaging reduces the variance of the measured intensity, which approaches the averaged speckle intensity $I(x) \to E[I^o(x)]$.

In general the mean of the measured intensity (averaged over spatially neighboring pixels) is not affected much by the pixel size. However, the intensity variance (or the speckles variance), is reduced as we increase the pixel size. The speckle variance is roughly

$$\left(\frac{\Delta_\Phi}{\Delta_x}\right)^2 E[I^o(x)] \tag{46}$$

since a single 2D pixel averages $n = (\Delta_x/\Delta_\Phi)^2$ diffraction blur windows, this argument explains why increasing the aperture width can reduce, or eliminate speckles: when the aperture is wide the diffraction blur Δ_Φ is small and hence the ratio Δ_x/Δ_Φ increase. Thus a single pixel Δ_x averages more independent intensities variables $I^o(x)$ and the variance due to speckles is reduced.

Another useful way to reduce speckles which will be discussed in Section A.2.1, is to use wide angle illumination.

Summary. The main source of speckles is that the complex *amplitude* of the field is averaged over a local diffraction blur window. Since for rough surfaces the phase varies rapidly, the mean amplitude is a random variable whose magnitude vary drastically. To eliminate speckles we need to average the *intensity* of a sufficient

number of independent mean amplitude variables. The size of the diffraction blur window can be reduced using a wider aperture.

As we see in the next subsection in optical coherent tomography the measured quantity is the *amplitude* of the signal rather than the *intensity*. This amplitude is averaged over the entire pixel area, however if the pixel area is wider than the diffraction blur the contrast of the signal we hope to measure is reduced. Therefore we cannot image a speckle-free signal (using a wide numerical aperture). To remove speckles we must averaged the intensities of several pixels in a computational post processing.

A.1.2 Speckles in Light Transport Interferometry

We start by rewriting the intensity measured by our setup while taking into account diffraction blur. Adding diffraction to Equation (28), the intensity arriving at position x is

$$I^o_{i,\tau}(x) = \left\langle |W_{\Delta_\Phi} * (u_s(x,t) + u_{r,\tau}(x,t))|^2 \right\rangle_t \tag{47}$$

and after averaging over the pixel area we measure:

$$I_{i,\tau}(x) = \Pi_{\Delta_x} * I^o_{i,\tau}(x) \tag{48}$$

Let us consider the intensity from a camera focused at plane $z = z_0$. As we discussed in Section 4, we can decompose the measured intensity as the intensity of the scattered and reference waves plus the correlation between them,

$$\begin{aligned} I^o_{i,\tau}(x) &= \left\langle |W_{\Delta_\Phi} * u_s(x,t)|^2 \right\rangle_t && (49)\\ &+ \left\langle |W_{\Delta_\Phi} * u_{r,\tau}(x,t)|^2 \right\rangle_t \\ &+ 2Re\left\langle (W_{\Delta_\Phi} * u_s(x,t))^* \cdot (W_{\Delta_\Phi} * u_{r,\tau}(x,t)) \right\rangle_t \\ &= I^o_s(x) + I^o_r(x) && (50)\\ &+ 2Re\left\langle (W_{\Delta_\Phi} * u_s(x,t))^* \cdot u_{r,\tau}(x,t)) \right\rangle_t \\ &= I^o_s(x) + I^o_r(x) + 2Re\left(W_{\Delta_\Phi} * C_\tau(x)\right) && (51) \end{aligned}$$

where the transformation from Equation (49) to Equation (50) used the fact that for coaxial plane illumination $u_{r,\tau}(x,t)$ is constant at x hence $W_{\Delta_\Phi} * u_{r,\tau}(x,t) = u_{r,\tau}(x,t)$.

In a similar way the intensity collected over the area of a sensor pixel is:

$$I_{i,\tau} = \Pi_{\Delta_x} * \left(I^o_s + I^o_{r,\tau}\right) + 2Re(\Pi_{\Delta_x} * W_{\Delta_\Phi} * C_\tau) \tag{52}$$

In our system, the intensity components are measured separately and subtracted and the measurement of interest is,

$$\begin{aligned} T^m(x,\tau) &\triangleq I_{i,\tau}(x) - (I_s(x) + I_{r,\tau}(x)) \\ &= 2Re(\Pi_{\Delta_x} * W_{\Delta_\Phi} * C_\tau) \end{aligned} \tag{53}$$

Equation (53) reveals that taking into account the finite pixel and aperture size, we eventually measure a blurred version of the correlation function.

The fact that we eventually measure *amplitude* rather than intensity is crucial for understanding speckle statistics. As discussed previously (in particular, Equation (45)), in standard imaging the amplitude is averaged over the diffraction blur window W_{Δ_Φ} of width Δ_Φ, and the *intensity* (power magnitude) of the diffraction windows W_{Δ_Φ} is further blurred by the pixel area. The intensity $I^o(x)$ of individual Δ_Φ-sized windows can have large variation leading to noisy speckle structure. However, if the diffraction blur width is smaller than the pixel area $\Delta_\Phi \ll \Delta_x$, speckle variation is reduced when integrating over the pixel area. This reduction in speckle variation does not happen in our measurements, since integrating over the pixel area Δ_x *sums amplitudes rather than intensities*. If $\Delta_\Phi \ll \Delta_x$, blur over the pixel area reduces the interference contrast and attenuates the information we aim to extract.

To remove speckles we need to further blur the intensity of the measured data in a computational post processing and display

$$\hat{T}^m(x,\tau) = \sqrt{g * |T^m(x,\tau)|^2} \tag{54}$$

where g blurs over the τ or x dimensions (or both). The exact choice of g can be application dependent and determined by the desired tradoffs between spatial and temporal resolutions. This blurring post-processing reduces resolution. For example given a fixed aperture size, to achieve the same resolution as in standard imaging we have to capture the data in a high magnification factor m, such that the pixel size Δ_x will be smaller than the diffraction width Δ_Φ. The major penalty for such larger magnification is reduction in depth of field and filed of view.

A.2 Speckles Under Area Light Sources

As mentioned above, in standard imaging the use of area light sources can largely reduce speckle variation. We review how this is achieved, and then explain why this solution does not carry to the light transport interferometry framework.

For laser sources it is possible to generate very strong amount of energy with a very narrow source which can be reasonably modeled as an ideal pinhole light source. White light sources, by contrast, are usually weaker and to generate high energy multiple nearby emitters are combined to form an area light source. While illumination emitted by pinhole laser sources is spatially coherent, due to the multiple emitters, area light sources are spatially incoherent. To understand the effect of area light sources we now also need to discuss spatial coherence. We start with conventional imaging, and then discuss the imaging for interferometry.

A.2.1 Speckles in Standard Imaging

As mentioned before, the overall field from the extended source is the superposition of directional fields

$$u(x,t) = \int_{\theta_0-\Theta}^{\theta_0+\Theta} u_\theta(x,t) d\theta \tag{55}$$

Since the directional sources are independent, there is zero correlation between the fields from different directions and hence the intensity of the extended source is the mean intensity from all directional sources:

$$I^o(x) = \left\langle |u(x,t)|^2 \right\rangle_t = \int_{\theta_0-\Theta}^{\theta_0+\Theta} \left\langle |u_\theta(x,t)|^2 \right\rangle d\theta = \int_{\theta_0-\Theta}^{\theta_0+\Theta} I^o_\theta(x) d\theta \tag{56}$$

Similarly

$$I(x) = \int_{\theta_0-\Theta}^{\theta_0+\Theta} I_\theta(x) d\theta \tag{57}$$

The fact that the measured intensity $I(x)$ is the mean of directional intensities I_θ implies a further reduction in speckle contrast, as we are averaging more independent variables. It can be shown that using a wider illumination effectively reduces the width of the diffraction blur.

When the scene is illuminated by a plane wave at incident angle θ, the field reaching the z_o plane is

$$u_{in,\theta}(x,t) = u_{in,\theta}\left(0, 0, t - \frac{\sin\theta \cdot x + \cos\theta \cdot z_o}{c}\right) \tag{58}$$

For standard imaging the recorded intensity is the averaged intensity of all wavelengths in the source and the spectral width of the source does not impose special interference. Therefore, in the coming lines we will focus the discussion at the behavior of monochromatic sources and ignore spectral bandwidth. Assuming illumina-

tion angles θ are small, u_{in} at the plane z_o is simply a sinusoid of frequency θ/λ, which up to multiplicative constants equals

$$u_{in,\theta}(x,t) = e^{\frac{-2\pi i}{\lambda}(\theta \cdot x - c \cdot t)} \tag{59}$$

The incident field interacts with the scene, reflecting and scattering back toward the $z = z_o$ plane. Using the notation of Section 4, the relationship between the ingoing and outgoing fields is the sum of contributions of all paths in the scene. When the light travels a path of length τ it achieves a time delay of τ/c and accordingly its phase varies by $e^{ik\tau}$. We denote by $\mathfrak{a}(x)$ the contribution of all paths ending at x

$$\mathfrak{a}(x) = \int_\tau e^{ik\tau} \int_\xi \mathcal{T}(x+\xi, x, \tau) \tag{60}$$

The outgoing field at direction θ is simply the multiplication of $\mathfrak{a}(x)$ by the incoming field:

$$u_\theta(x,t) \quad = \quad u_{in,\theta}(x,t) \cdot \mathfrak{a}(x) = e^{\frac{-2\pi i}{\lambda}(\theta \cdot x - c \cdot t)} \mathfrak{a}(x) \tag{61}$$

We want to show that wide illumination can significantly reduce speckle contrast. The following claim states mathematically the fact that if the illumination angle is significantly larger than the numerical aperture the measured intensity is the field blurred with a sinc whose width is inversely proportional to the illumination angle Θ rather than the aperture angle Φ.

Claim 4. *Let*

$$\Delta_\Theta = \frac{\lambda}{2\Theta} \tag{62}$$

If $\Delta_\Phi \gg \Delta_\Theta$ *the measured intensity is approximately*

$$I(x) \approx \Pi_{\Delta_x} * W_{\Delta_\Phi} * I^p(x), \quad I^p(x) = \left\langle |W_{\Delta_\Theta} * u(x,t)|^2 \right\rangle_t \tag{63}$$

Proof: Let us denote by $\mathfrak{A}(w)$ the Fourier transform of the surface modulation function $\mathfrak{a}(x)$ over the window $W_{\Delta_\Phi}(x)$:

$$\mathfrak{A}(\omega) = \int \mathfrak{a}(x) \cdot W_{\Delta_\Phi}(x) \cdot e^{2\pi i(\omega \cdot x)} dx \tag{64}$$

Using Equations (61) and (45), we can see that $I^o(x)$ is essentially the Fourier power spectrum of the surface modulation function $\mathfrak{a}(x)$ at frequency θ/λ and $I^o_\theta = |\mathfrak{A}(\theta/\lambda)|^2$. The intensity of the extended source is the power spectrum of $\mathfrak{A}(\omega)$, integrated over all frequencies in the range $\lambda^{-1}[\Theta, \Theta]$, or equivalently the norm of $\mathfrak{A}(\omega)$ multiplied by a rect filter of width $2\Theta/\lambda$:

$$I^o(x) = \int_{\theta_0 - \Theta}^{\theta_0 + \Theta} I^o_\theta(x) d\theta = \int \left| \mathfrak{A}(\omega) \cdot \Pi_{2\Theta/\lambda}(\omega) \right|^2 d\omega \tag{65}$$

The Fourier transform of a rect filter $\Pi_{2\Theta/\lambda}$ of width $2\Theta/\lambda$ is a sinc.

$$W_{\Delta_\Theta}(x) = \mathrm{sinc}\left(\frac{x}{\Delta_\Theta}\right) \qquad \text{with } \Delta_\Theta = \frac{\lambda}{2\Theta}. \tag{66}$$

From the convolution theorem and Parsevall's theorem the norm of the signal in Equation (65) can also be computed in the primal domain, as the norm of $\mathfrak{a}(x) \cdot W_{\Delta_\Phi}$, convolved with W_{Δ_Θ}:

$$I^o(x_0) = |W_{\Delta_\Theta} * (\mathfrak{a}(x) \cdot W_{\Delta_\Phi}(x - x_0))|^2 \tag{67}$$

Assuming that $\Delta_\Phi \gg \Delta_\Theta$ the windowing with W_{Δ_Φ} is mostly effective in setting the range and

$$I^o \approx W_{\Delta_\Phi} * |W_{\Delta_\Theta} * \mathfrak{a}|^2 \tag{68}$$

Note that $W_{\Delta_\Theta} \cdot \mathfrak{a} = W_{\Delta_\Theta} \cdot u_0$ (where u_o denotes the field at $\theta = 0$, thus

$$I^o \approx W_{\Delta_\Phi} * |W_{\Delta_\Theta} * u_0|^2 \tag{69}$$

combining Equations (69) and (45), we obtain the desired relation,

$$I \approx \Pi_{\Delta_x} * W_{\Delta_\Phi} * |W_{\Delta_\Theta} * u_0|^2 \tag{70}$$

□

A.2.2 Speckles in Light Transport Interferometry

To compute the intensity measured by our system under incoherent illumination we integrate Equation (49) over all illumination directions:

$$\begin{aligned} {I^o}_{i,\tau}(x) = \Pi_{\Delta_x} * \int_\theta & \left\langle |W_{\Delta_\Phi} * u_{s,\theta}(x,t)|^2 \right\rangle_t \\ + & \left\langle |W_{\Delta_\Phi} * u_{r,\theta,\tau}(x,t)|^2 \right\rangle_t \\ + & 2Re \left\langle (W_{\Delta_\Phi} * u_{s,\theta}(x,t))^* \cdot (W_{\Delta_\Phi} * u_{r,\theta,\tau}(x,t)) \right\rangle_t \end{aligned} \tag{71}$$

To study this let's consider the term $W_{\Delta_\Phi} * u_{r,\theta,\tau}(x,t)$. Assuming the reference arm is a flat mirror u_r is a sum of sinusoids. These sinusoids are convolved with a sinc, which corresponds in the frequency domain to an impulse multiplied by a rect filter. Thus $u_{r,\tau}$ convolved with a sinc is zero or unaffected, depending of the frequency of the sinusoid:

$$W_{\Delta_\Phi} * u_{r,\theta,\tau}(x,t) = \begin{cases} u_{r,\theta,\tau}(x,t) & \theta < \Phi \\ 0 & \theta > \Phi \end{cases} \tag{72}$$

This can also be understood intuitively: the mirror in the reference arm simply reflects the incoming illumination to the camera. If the illumination angle is larger than the acceptance angle of the aperture ($\theta > \Phi$), it is cropped by the aperture.

Therefore, the correlation component of the interference signal in Equation (71) reduces to

$$\Pi_{\Delta_x} * W_{\Delta_\Phi} * \int_\theta \left\langle u_{s,\theta}(x,t)^* \cdot u_{r,\theta,\tau}(x,t) \right\rangle_t \tag{73}$$

Using Equation (34) this reduces to

$$(\Pi_{\Delta_x} * W_{\Delta_\Phi} * W_{\Delta_c}) * (e^{i\bar{k}\tau} G_{\Delta_k^{-1}}(\tau)) * \mathcal{T}(x,x,\tau). \tag{74}$$

Equation (74) implies that the interference signal we will measure is the desired pathlength-resolved light transport func- tion $\mathcal{T}$, blurred spatially with $\Pi_{\Delta_x} * W_{\Delta_\Phi} * W_{\Delta_\Theta}$ and blurred in time with the temporal coherence point spread function $e^{i\bar{k}\tau} e^{-\frac{1}{2}(\Delta_k(\tau))^2}$. As discussed above, this blurring is applied to the amplitude rather than the intensity, therefore it can reduce the contrast of the interference signal, unless the imaging parameters are selected such that the blur $\Pi_{\Delta_x} * W_{\Delta_\Phi}$ is negligible with respect to Δ_Θ.

B Interference Energy Reduction

When interferometry is applied on a dense scattering media, neighbooring particles emit multiple scattered fields. In Section 4 we have derived the correlation component of the recorded interference signal and showed that it is a blurred version of the ideal pathlength-

resolved light transport function function $\mathcal{T}$,

$$C_\tau(x) \propto e^{-i\bar{k}\tau} \iint \mathcal{T}(x+\xi, x, \zeta) e^{i\bar{k}\zeta} e^{-\frac{1}{2}(\Delta_k(\zeta-\tau))^2} d\zeta d\xi \quad (75)$$

When the particle distance is smaller than the wavelength, the interference signal (the blurred $\mathcal{T}$ function in Equation (32)) averages scalars with different phases. As a result, interference between neighboring points will attenuate the intensity of the measured signal. One may be concerned that this attenuation will destroy all information we hope to capture. To address this concern, in this section we show that the exact same interference phenomenon occurs in the intensity image I_s that would be captured using a conventional camera, and therefore that the power of the correlation component of the interference signal is proportional to that of the standard intensity image, with no extra attenuation involved.

While interference between particles does occur, the particle locations are sufficiently random so that the interference is not perfectly destructive. The interference between neighboring paths is the same natural phenomenon happening when measuring a standard intensity image. As most mediums have a sufficiently random structure, the measured intensity image via scattering media contains nonzero energy. We quantify this argument below.

To simplify notation, we assume here coherent illumination and neglect integration over illumination angles. Since in this case the wave is spatially coherent, we shorten notation and write

$$\mathcal{T}^c(x,\tau) = \int \mathcal{T}(\xi, x, \tau) d\xi. \quad (76)$$

The intensity of the scattered field. Recall that the measured intensity is $I_s(x) =< |u_s(x,t)|^2 >_t$. We can write

$$\begin{aligned}
&\langle u_s(x,t)^*, u_s(x,t)\rangle_t \\
&= \int_t \left(\int_{\tau_1} \mathcal{T}^c(x,\tau_1)^* \int_{k_1} a_{k_1} \cdot e^{-ik_1(ct-\tau_1)} \right) \cdot \\
&\quad \left(\int_{\tau_2} \mathcal{T}^c(x,\tau_2) \int_{k_2} a_{k_2} \cdot e^{ik_2(ct-\tau_2)} \right) \\
&= \iint_{\tau_1,\tau_2} \mathcal{T}^c(x,\tau_1)^* \mathcal{T}^c(x,\tau_2) \int_k |a_k|^2 \cdot e^{ik(\tau_2-\tau_1)} \\
&= \iint_{\tau_1,\tau_2} \mathcal{T}^c(x,\tau_1)^* \mathcal{T}^c(x,\tau_2) e^{i\bar{k}(\tau_2-\tau_1)} g_{\Delta_k^{-2}}(\tau_2-\tau_1) \\
&= \int_{\tau_1} \mathcal{T}^c(x,\tau_1)^* e^{-i\bar{k}\tau_1} \int_\epsilon \mathcal{T}^c(x,\tau_1+\epsilon) e^{i\bar{k}(\tau_1+\epsilon)} G_{\Delta_k^{-2}}(\epsilon)
\end{aligned} \quad (77)$$

where we use $G_{\sigma^2}(\epsilon)$ as a shorten notation for a gaussian function with variance σ^2:

$$G_{\sigma^2}(\epsilon) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\epsilon)^2}{2\sigma^2}} \quad (78)$$

To express Equation (77) let us define $\tilde{\mathcal{T}^c}_{\sigma^2}$ as the signal $\mathcal{T}^c$ times a sinusoid, blurred over a window of width σ

$$\tilde{\mathcal{T}^c}_{\sigma^2}(x,\tau) = (\mathcal{T}^c(x,\tau)e^{i\bar{k}\tau}) * G_{\sigma^2} = \int \mathcal{T}^c(x,\tau+\epsilon) e^{i\bar{k}(\tau+\epsilon)} G_{\sigma^2}(\epsilon) d\epsilon \quad (79)$$

We can then express Equation (77) as an inner product between b and its blurred version

$$\langle u_s(x,t)^*, u_s(x,t)\rangle_t = \left\langle \mathcal{T}^c(x,\tau)^* e^{-i\bar{k}\tau}, \tilde{\mathcal{T}^c}_{\Delta_k^{-2}}(x,\tau) \right\rangle_\tau \quad (80)$$

Equivalently we can think of a blur with a smaller support on both sides

$$\langle u_s(x,t)^*, u_s(x,t)\rangle_t = \left\langle \tilde{\mathcal{T}^c}_{\Delta_k^{-2}/2}(x,\tau)^*, \tilde{\mathcal{T}^c}_{\Delta_k^{-2}/2}(x,\tau) \right\rangle_\tau \quad (81)$$

This holds since

$$G_{\Delta_k^{-2}/2} * G_{\Delta_k^{-2}/2} = G_{\Delta_k^{-2}} \quad (82)$$

(to prove that relationship we can express the convolution in the frequency domain and use the convolution theorem).

In short the intensity of the scattered field equals the magnitude of a blurred version of the pathlength-resolved light transport function $\mathcal{T}^c(x,\tau)$. Of course, since epsilon shifts in the path length result in different phases, interference takes place and reduces the intensity of the signal. However, given that when we image scattering materials we measure some intensity, the implication is that we do not have perfectly destructive interference.

The intensity of the correlation signal. As derived in Section 4 the correlation component of the OCT signal is also a blurred version of the pathlength-resolved light transport function $\mathcal{T}^c$, which can be written as:

$$\langle u_{r,\tau_o}(x,t)^*, u_s(x,t)\rangle_t = 2Re\left(e^{-i\bar{k}\tau_o} \cdot \tilde{\mathcal{T}^c}_{\Delta_k^{-2}}(x,\tau_o) \right) \quad (83)$$

That is, the interference signal will correspond to the pathlength-resolved light transport function, blurred with a window of width Δ_k^{-2}. This blurring reduces contrast since the signal $\mathcal{T}^c(x,\tau)$ has changing phase, but it does not wash out the magnitude completely, as it does not wash the intensity of the scattered field itself. As seen in Equation (81), the scattered field is blurred by a slightly narrower window of size $\Delta_k^{-2}/2$, so we can expect that the magnitude of the interference signal will be somewhat weaker. The interference signal will correspond to the pathlength decomposition of a signal with a slightly longer coherence length, but changing the coherence length does not completely wash out the intensity of the scattered wave.

C Acquisition setup

In this section, we provide details about our implementation of the setup shown in Figure 8. For convenience, the schematic is replicated in Figure 19, along with a photograph of our setup. All references to parts are with respect to the schematic of Figure 19.

Light source. We use three different light sources (part (a)), depending on the application. The first source is a high-power white LED, with an output power of $9\,\mathrm{W}$, which we use in combination with color filters. The second is a red ($625\,\mathrm{nm}$ central wavelength) LED source of output $0.7\,\mathrm{W}$. Even though it has overall less power, all of its output is concentrated within a narrow band of $25\,\mathrm{nm}$, resulting in an overall higher intensity and analogous coherence length compared to the white LED and color filter combination. Additionally, the output of the red LED has a much smaller numerical aperture, making collimation easier and improving interference contrast. On the downside, it is not possible to do RGB scans with this source. We used the red LED to obtain the scattering measurements of Section 6, for which contrast is critical. When using either of these two LED sources, immediately in front of the LED we place an adjustable aperture (part (b)), which we use to control the angular extent of the source.

The third source is an SLD of output power $5\,\mathrm{mW}$, central wavelength $650\,\mathrm{nm}$, and bandwidth $20\,\mathrm{nm}$. We found it necessary to use a spatial filter to clean-up the output of the raw diode, which contained undesirable angular components. When using the SLD (with the spatial filter), we remove the adjustable aperture, as the SLD is already practically a pinhole source.

Illumination lens. We use a $200\,\mathrm{mm}$ prime compound lens (part

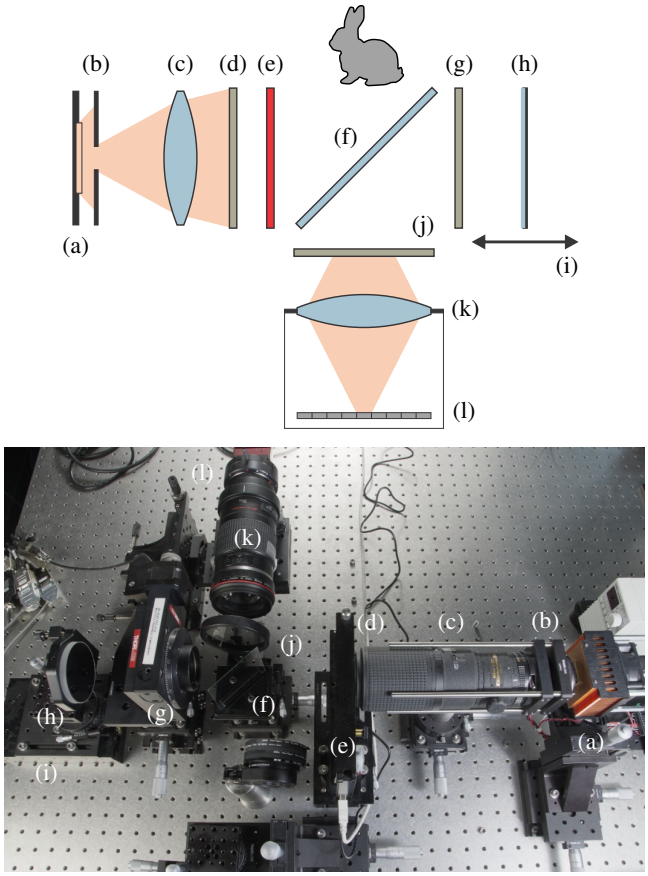


Figure 19: *Schematic (top) and physical implementation (bottom) of our optical setup.*

(c)) to collimate the output of our sources. We also experimented with off-the-shelf, AR-coated achromatic doublets. The compound lens unsurprisingly performed better in terms of reducing chromatic aberration, improving collimation, and additionally slightly improving light efficiency. However, the most critical improvement and the reason why we used the compound lens was that it drastically reduced interreflections compared to the off-the-shelf lens, despite the latter's coating. We discuss interreflections later. Finally, we only experimented with telephoto-range focal lengths, as beam collimation is critical for interference contrast. Off-the-shelf $100\,\mathrm{mm}$ lenses resulted in considerably worse SNR compared to off-the-shelf $200\,\mathrm{mm}$ lenses. For this reason, we selected the latter, despite the better light efficiency of lenses with shorter focal lengths.

Beamsplitter. Beamsplitters typically come in three types: cube, pellicle, and plate. Pellicle beamsplitters exhibit no chromatic aberration, and are also free of secondary reflections. However, because they are made of a very thin membrane, they are extremely sensitive to vibrations arising from noise or other environment sources. We discuss vibrations later. Cube beamsplitters are rigid, but create many secondary reflections and have large chromatic aberration. We experimented with both types, and in both cases their respective limitations resulted in severely reduced detection SNR. As a compromise between the two, we use a thin plate beamsplitter (part (f)).

Mirrors. We use high-quality first surface mirrors of $\lambda/8$ flatness (part (h)), though in practice we found no difference in performance when using lower-quality mirrors. For diagonal probing (Figure 6(a)), we use a single, circular mirror that can be mounted rigidly, to reduce vibrations. For anti-diagonal probing (Figure 6(b)), we use a pair of square mirrors, mounted to form a right angle. For anti-diagonal probing of scenes that are not two-dimensional, we use a monolithic hollow mirror retroreflector (three mirror surfaces forming a cube corner). We use extra square mirrors for the alignment procedure described below.

Translation stage. We use a high-end translation stage (l) with minimum incremental motion of $10\,\mathrm{nm}$ and low-noise operation. This is necessary in order to minimize surface vibrations induced by the stage's motion. We discuss this in more detail later.

Camera lens. Given the size of our scenes, achieving at least $1:1$ (real size) magnification is a key specification for the imaging lens. We image using a $180\,\mathrm{mm}$ prime compound macrolens (part (k)). The long focal length is necessary in order to have enough working distance, even at large magnifications, for all the optical parts between the camera and the target scene or reference mirror (beamsplitter, polarizers, color filters, and optical mounts). Additionally, the long focal length improves interference contrast.

Camera. We use a camera with a high sensitivity CCD sensor (part (l)). The $8\,\mathrm{MP}$ sensor has a pixel pitch of $4\,\mu\mathrm{m}$. As we discuss in Section 5, small pixel pitch is an important factor for improving interference contrast. The large sensor size is important for achieving sufficient fields of view, especially at large magnifications. Finally, given the contrast-exposure time trade-off described in Figure 5, the high sensitivity of the sensor is critical for keeping capture times reasonable.

Color filters. We use hard-coated bandpass color filters (part (e), as their high transmission is helpful for reducing capture times. One downside is that the reflective coating of these filters creates a very strong interreflection that can negatively affect interference contrast. We discuss this later. We use filters of different bandwidths, $3\,\mathrm{nm}$, $10\,\mathrm{nm}$, $25\,\mathrm{nm}$, depending on the desired pathlength resolution (Figure 4). To capture measurements at multiple wavelengths, we use filters centered at different wavelengths. The exact center wavelengths change slightly for filters of different bandwidths; for the $25\,\mathrm{nm}$ filters we use in most experiments, they are: $625\,\mathrm{nm}$ (red), $525\,\mathrm{nm}$ (green), and $450\,\mathrm{nm}$ (blue). The center wavelengths We use a motorized filter wheel to automatically change filters. This is necessary in order to repeatably perform multiple scans.

Polarizers. We use a high-quality wire grid polarizer at the source side (part (d)). Such polarizers have high parallel transmission, which helps reduce capture times. We use simple, laminated film polarizers at the camera and reference arms (parts (g) and (j)), as those polarizers need to be at least as big as the desirable maximum field of view—wire-grid polarizers only come in small sizes. The camera and mirror polarizers are mounted on motorized rotation stages, so that they can be rotated (for instance, to change from parallel to cross-polarization or to do HDR) without user intervention. As with the color filters, it is critical to be able to reconfigure the polarizer positions automatically and repeatably when performing multiple scans.

Neutral density filters. We use absorptive neutral density filters. Reflective filters have strong interreflections, which can severely reduce interference contrast.

Interreflections. In order to maximize interference contrast, and therefore SNR, it is important to minimize interreflections between the optics of our setup. These interreflections add a constant intensity component to the captured images, which does not interfere at any position of the reference arm. As a result, a portion of the camera's dynamic range is used up just by interreflections, suppressing the interference signal and directly reducing the SNR of our measurements. In practice, we found that minimizing interreflections is critical for obtaining useful interferometric measurements in all but the most trivial scenes (mirror or single diffuse planes).

For this reason, we made sure that all of our optical components had multi-layer anti-reflective coating. We used a third-party coating

company for components that were not available in coated versions by their providers (such as the neutral density filters).

In practice, we also found it helpful to slightly tilt certain optical components, so that any interreflections caused by them would be directed away from the optical path. This was most important for the color filters, due to the large reflections they created, but it also applied to the camera and mirror polarizers.

Vibrations. Another factor that can critically reduce interference contrast and SNR is vibrations of optical components. The motion from such vibrations blurs speckle noise during exposure, resulting in decreased contrast. The decrease is exponential with respect to the vibration amplitude. As a result, even submicron vibrations can effectively remove all interference contrast for scenes that require exposure times of more than a few milliseconds.

For this reason, we built our setup on an optical table with a passively damped table top and a pneumatic isolation system, to reduce surface and environment vibrations respectively. Moreover, we placed an enclosure system around the setup during acquisition, to reduce vibrations from environment sources such as acoustic noise or air flow.

Within the setup, a major source of mechanical vibrations is the translation stage at the reference arm. We experimented with DC stepper motor and piezo-electric stages, as well as a high-end stage designed for nanometer resolutions. The very low-noise operation of the latter significantly improved performance compared to the other two categories, though the piezo-electric stages also gave good measurement contrast.

Finally, rigid and stable mechanical mounts for all optical components can significantly improve interference contrast. We used a laser Doppler vibrometer to identify parts that were vibrating during the scanning process, and improve mounting as necessary.

Motion control. The optical setup requires careful alignment (described later) of the various optical components at accuracies of a few microns. To achieve this, we mount most components on manual micro-positioning stages, with various degrees of freedom as necessary. The reference mirror, camera, beamsplitter and source LED have mounts providing five degrees of freedom (XYZ translation, tilt, and rotation), whereas the illumination lens has three degrees of freedom (XY translation and rotation).

Alignment. The optical setup requires aligning the optical axis of the illumination, reference mirror, and camera (through the beamsplitter). We perform this alignment as described below.

First, we mount a pinhole in front of the illumination lens, to reduce lighting to a narrow beam. In practice, we have an adjustable aperture permanently mounted at the front of the lens, and close it for alignment.

Second, we mount a mirror between the beamsplitter and the illumination lens, at an angle of 45° relative to the beamsplitter (there exist carefully machined beamsplitter mounts that enable this). We rotate and tilt the beamsplitter mount, until the beam is reflected by the mirror back at the center of the pinhole. This ensures that the beamsplitter is aligned at 45° relative to the illumination direction.

Third, we remove the mirror, leaving the pinhole in place. The beam now travels through the beamsplitter towards the reference mirror, where it is reflected. To align the reference mirror with the optical axis, we rotate and tilt the reference mirror until the reflected beam returns at the center of the pinhole.

Fourth, we remove the pinhole from the illumination lens, and focus the camera lens at infinity. The infinity focus means that the illumination axis will be parallel with the camera's optical axis when all of the collimated light is imaged at the center of the camera sensor. We can make this happen by rotating and tilting the camera while looking at a live stream of the sensor. Once the two axes are parallel, we refocus the camera lens and translate the camera without changing its direction, in order to spatially center the collimated beam relative to the sensor.

Finally, in many cases it is useful to place another aligned mirror at the targer arm, as in the basic Michelson interferometer setup of Figure 2. This can be used to measure resolution, through the experiment described in Figure 4, or to fine-tune various imaging parameters for maximum interference contrast. The alignment can be done by first focusing the camera lens at infinity, then tilting and rotating the target mirror until its reflection is imaged at the center of the camera sensor.

Component list. In Table 1, we provide a list of the major components used in our setup. We omit standard optical mounts and micropositioning stages, which can be found in most optical equipment providers.

References

ABRAMSON, N. 1983. Light-in-flight recording: high-speed holographic motion pictures of ultrafast phenomena. *Applied Optics 22*, 2 (Jan), 215–232.

BAI, J., CHANDRAKER, M., NG, T.-T., AND RAMAMOORTHI, R. 2010. A Dual Theory of Inverse and Forward Light Transport. In *Proceedings of the 11th European Conference on Computer Vision*, ECCV'10, 294–307.

COSSAIRT, O., MATSUDA, N., AND GUPTA, M. 2014. Digital refocusing with incoherent holography. In *Computational Photography (ICCP), 2014 IEEE International Conference on*, 1–9.

DORRINGTON, A. A., GODBAZ, J. P., CREE, M. J., PAYNE, A. D., AND STREETER, L. V. 2011. Separating true range measurements from multi-path and scattering interference in commercial range cameras. In *Proc. SPIE*, vol. 7864.

GOODMAN, J. W. 1968. *Introduction to Fourier Optics*. McGraw-Hill Book Company.

GOODMAN, J. W. 2000. *Statistical Optics*. Wiley Classics Library.

GUPTA, M. AND AGRAWAL, A. AND VEERARAGHAVAN, A. AND NARASIMHAN, S.G. 2011. Structured light 3d scanning in the presence of global illumination. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, 713–720.

HARIHARAN, P. 2003. *Optical interferometry*. Elsevier.

HEIDE, F., HULLIN, M. B., GREGSON, J., AND HEIDRICH, W. 2013. Low-budget Transient Imaging Using Photonic Mixer Devices. *ACM Trans. Graph. 32*, 4 (July), 45:1–45:10.

HEIDE, F., XIAO, L., KOLB, A., HULLIN, M. B., AND HEIDRICH, W. 2014. Imaging in scattering media using correlation image sensors and sparse convolutional coding. *Optics Express 22*, 21 (Oct), 26338–26350.

HUANG, D., SWANSON, E., LIN, C., SCHUMAN, J., STINSON, W., CHANG, W., HEE, M., FLOTTE, T., GREGORY, K., PULIAFITO, C., AND FUJIMOTO, G. 1991. Optical coherence tomography. *Science 254*, 5035, 1178–1181.

JARABO, A., MARCO, J., MUÑOZ, A., BUISAN, R., JAROSZ, W., AND GUTIERREZ, D. 2014. A Framework for Transient Rendering. *ACM Trans. Graph. 33*, 6 (Nov.), 177:1–177:10.

KADAMBI, A., WHYTE, R., BHANDARI, A., STREETER, L., BARSI, C., DORRINGTON, A., AND RASKAR, R. 2013. Coded Time of Flight Cameras: Sparse Deconvolution to Address Multipath Interference and Recover Time Profiles. *ACM Trans. Graph. 32*, 6 (Nov.), 167:1–167:10.

LEVIN, A., GLASNER, D., XIONG, Y., DURAND, F., FREEMAN, W., MATUSIK, W., AND ZICKLER, T. 2013. Fabricating

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
2 inch hard coated bandpass filter, 488 nm CWL, 3 nm BW	1	68-904	Edmund Optics
2 inch hard coated bandpass filter, 532 nm CWL, 3 nm BW	1	68-907	Edmund Optics
2 inch hard coated bandpass filter, 635 nm CWL, 3 nm BW	1	68-911	Edmund Optics
2 inch hard coated bandpass filter, 450 nm CWL, 10 nm BW	1	65-201	Edmund Optics
2 inch hard coated bandpass filter, 532 nm CWL, 10 nm BW	1	65-216	Edmund Optics
2 inch hard coated bandpass filter, 635 nm CWL, 10 nm BW	1	65-228	Edmund Optics
2 inch hard coated bandpass filter, 450 nm CWL, 25 nm BW	1	86-665	Edmund Optics
2 inch hard coated bandpass filter, 525 nm CWL, 25 nm BW	1	87-801	Edmund Optics
2 inch hard coated bandpass filter, 625 nm CWL, 25 nm BW	1	87-803	Edmund Optics
3 inch plate beamsplitter	1	46-642	Edmund Optics
motorized filter wheel	1	84-889	Edmund Optics
3 inch square aluminum mirror	4	84-416	Edmund Optics
motorized filter wheel	1	84-889	Edmund Optics
colored SLD, 650 nm CWL, 10 nm BW, 5 mW power	1	EXS210035-02	EXALOS
LED development kit, cooler and controller	1	DK-136M-1	Luminus Devices
white LED, 9 W power	1	CBT-140-WCS-C15-UA120	Luminus Devices
broadband hollow retroreflector	1	UBBR2.5-1UV	Newport Corporation
ethernet driver for linear stage	1	XPS-Q2	Newport Corporation
large tilt and rotation platform	1	37	Newport Corporation
PWM drive module for linear stage	1	XPS-DRV02	Newport Corporation
ultra-precision linear motor stage, 2 inch travel	1	XMS100	Newport Corporation
200 mm compound lens	1	AF Micro Nikkor 200mm 1:4 D IF-ED	Nikon
controllable benchtop power supply	1	ZUP 20-20	TDK-Lambda
1 inch clear-aperture rotation stage	1	PRM1Z8	Thorlabs
1 inch wire grid linear polarizer	1	WP25M-VIS	Thorlabs
2 inch clear-aperture rotation stage	1	NR360S	Thorlabs
2 inch absorptive neutral density filter kit	1	NEK01S	Thorlabs
2 inch laminated film linear polarizer	1	LPVISE200-A	Thorlabs
3 inch laminated film linear polarizer	1	custom	Thorlabs
3 inch circular silver mirror	1	PF30-03-P01	Thorlabs
benchtop diode current controller	1	LDC205C	Thorlabs
colored LED, 625 nm CWL, 25 nm BW, 700 mW power	1	M625L3	Thorlabs
colored LED driver	1	LEDD1B	Thorlabs
linear translation stage, 2 inch travel	3	LT1	Thorlabs
micrometer-controlled linear translation stage, 1 inch travel	11	PT1	Thorlabs
micrometer-controlled rotation stage	8	PR01	Thorlabs
micrometer-controlled tilt platform	4	PY003	Thorlabs
servo motor controller for rotation stage	1	TDC001	Thorlabs
small lab jack	2	L200	Thorlabs
stepper motor controller for rotation stage	1	BSC201	Thorlabs
TE-cooled mount for TO-56 diodes	1	TCLDM9	Thorlabs
TEC diode controller	1	TTC001	Thorlabs

Table 1: *List of major components used in the optical setup of Figure 19.*

BRDFs at High Spatial Resolution Using Wave Optics. *ACM Trans. Graph. 32*, 4 (July), 144:1–144:14.

LUAN, X. 2001. Experimental investigation of photonic mixer device and development of TOF 3D ranging systems based on PMD technology. *PhD thesis, University of Siegen*.

NAYAR, S. K., KRISHNAN, G., GROSSBERG, M. D., AND RASKAR, R. 2006. Fast Separation of Direct and Global Components of a Scene Using High Frequency Illumination. *ACM Trans. Graph. 25*, 3 (July), 935–944.

NG, R., RAMAMOORTHI, R., AND HANRAHAN, P. 2003. All-frequency Shadows Using Non-linear Wavelet Lighting Approximation. *ACM Trans. Graph. 22*, 3 (July), 376–381.

O'TOOLE, M., AND KUTULAKOS, K. N. 2010. Optical Computing for Fast Light Transport Analysis. *ACM Trans. Graph. 29*, 6 (Dec.), 164:1–164:12.

O'TOOLE, M., RASKAR, R., AND KUTULAKOS, K. N. 2012. Primal-dual Coding to Probe Light Transport. *ACM Trans. Graph. 31*, 4 (July), 39:1–39:11.

O'TOOLE, M., MATHER, J., AND KUTULAKOS, K. 2014. 3D Shape and Indirect Appearance by Structured Light Transport. In *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*, 3246–3253.

O'TOOLE, M., HEIDE, F., XIAO, L., HULLIN, M. B., HEIDRICH, W., AND KUTULAKOS, K. N. 2014. Temporal Frequency Probing for 5D Transient Analysis of Global Light Transport. *ACM Trans. Graph. 33*, 4 (July), 87:1–87:11.

PEERS, P., MAHAJAN, D. K., LAMOND, B., GHOSH, A., MATUSIK, W., RAMAMOORTHI, R., AND DEBEVEC, P. 2009. Compressive Light Transport Sensing. *ACM Trans. Graph. 28*, 1 (Feb.), 3:1–3:18.

REDDY, D., RAMAMOORTHI, R., AND CURLESS, B. 2012. Frequency-space Decomposition and Acquisition of Light Transport Under Spatially Varying Illumination. In *Proceedings of the 12th European Conference on Computer Vision*, ECCV'12, 596–610.

SCHWARTE, R., HEINOL, H.-G., XU, Z., AND HARTMANN, K. 1995. New active 3D vision system based on rf-modulation interferometry of incoherent light. In *Proc. SPIE*, vol. 2588.

SEN, P., CHEN, B., GARG, G., MARSCHNER, S. R., HOROWITZ, M., LEVOY, M., AND LENSCH, H. P. A. 2005. Dual Photography. *ACM Trans. Graph. 24*, 3 (July), 745–755.

VELTEN, A., WU, D., JARABO, A., MASIA, B., BARSI, C., JOSHI, C., LAWSON, E., BAWENDI, M., GUTIERREZ, D., AND RASKAR, R. 2013. Femto-photography: Capturing and Visualizing the Propagation of Light. *ACM Trans. Graph. 32*, 4 (July), 44:1–44:8.

WANG, J., DONG, Y., TONG, X., LIN, Z., AND GUO, B. 2009. Kernel Nyström Method for Light Transport. *ACM Trans. Graph. 28*, 3 (July), 29:1–29:10.

WU, D., VELTEN, A., OTOOLE, M., MASIA, B., AGRAWAL, A., DAI, Q., AND RASKAR, R. 2014. Decomposing Global Light Transport Using Time of Flight Imaging. *International Journal of Computer Vision 107*, 2, 123–138.

WU, D., WETZSTEIN, G., BARSI, C., WILLWACHER, T., DAI, Q., AND RASKAR, R. 2014. Ultra-fast Lensless Computational Imaging through 5D Frequency Analysis of Time-resolved Light Transport. *International Journal of Computer Vision 110*, 2, 128–140.