Passive Micron-scale Time-of-Flight with Sunlight Interferometry

Alankar Kotwal[1], Anat Levin[2], and Ioannis Gkioulekas[1]

[1]Carnegie Mellon University, [2]Technion

Abstract

We introduce an interferometric technique for passive time-of-flight imaging and depth sensing at micrometer axial resolutions. Our technique uses a full-field Michelson interferometer, modified to use sunlight as the only light source. The large spectral bandwidth of sunlight makes it possible to acquire micrometer-resolution time-resolved scene responses, through a simple axial scanning operation. Additionally, the angular bandwidth of sunlight makes it possible to capture time-of-flight measurements insensitive to indirect illumination effects, such as interreflections and subsurface scattering. We build an experimental prototype that we operate outdoors, under direct sunlight, and in adverse environment conditions such as machine vibrations and vehicle traffic. We use this prototype to demonstrate, for the first time, passive imaging capabilities such as micrometer-scale depth sensing robust to indirect illumination, direct-only imaging, and imaging through diffusers.

1. Introduction

The recovery of 3D information for an imaged scene is one of the core problems of optics, imaging, computer vision, and other sciences. In particular, the ability to sense depth at axial resolutions of a few micrometers is of great importance for critical applications areas such as medicine, precision fabrication, material science, and robotics.

Existing contactless imaging techniques for micron-scale 3D sensing, such as interferometry and microscopy, require active illumination, most commonly from a coherent source. This makes these techniques impractical for use outdoors, in the presence of strong ambient illumination that overwhelms the active source, or in power-constrained applications. On the other hand, existing passive 3D sensing techniques, such as multi-view stereo, shape from shading, and depth from (de)focus, achieve resolutions of hundreds of micrometers at best, placing them out of scope for applications requiring micron-scale resolutions.

We change this state of affairs by introducing a completely passive micron-scale 3D sensing technique. Our technique is interferometric, and uses sunlight as the only light source, to capture full-frame time-of-flight measurements at axial resolutions of 5 micrometers. Our technique additionally takes advantage of the spatial incoherence of sunlight, to enable robust 3D sensing in the presence of severe indirect illumination effects (interreflections, subsur-
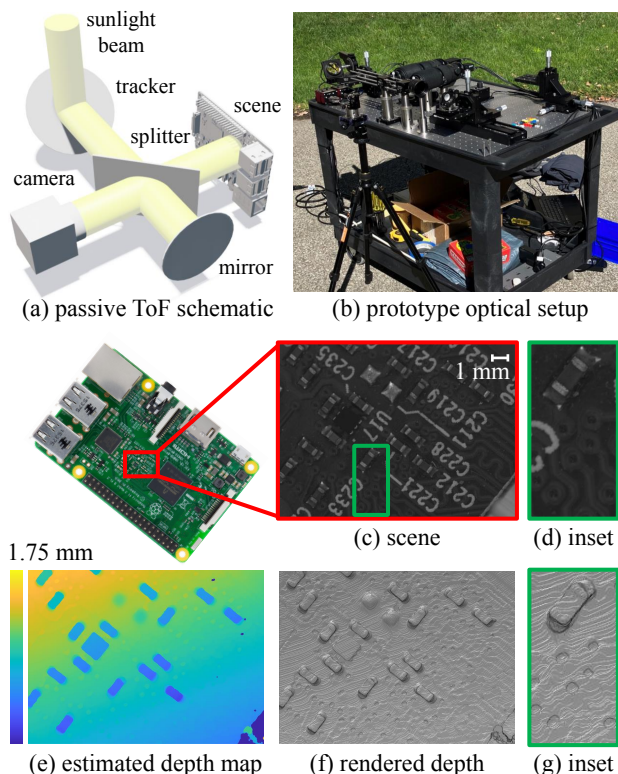

(a) passive ToF schematic   (b) prototype optical setup

1.75 mm

(c) scene   (d) inset

(e) estimated depth map   (f) rendered depth   (g) inset

Figure 1. **Using sunlight interferometry to passively reconstruct part of a circuit board.** **(a)** and **(b)** show a schematic and photograph of the sunlight interferometer we build for passive time-of-flight imaging. We use this system to reconstruct part of a Raspberry Pi circuit board that has multiple resistors, soldering pads, and tracks. **(c)** and **(d)** show a picture of the scene as seen through the imaging camera, along with an inset highlighting fine geometric features. **(e)** shows the estimated depth map, and **(f)** and **(g)** the corresponding rendered 3D surface. Our technique reconstructs fine features such as the PCB tracks and through-holes, despite operating outdoors under adverse environment conditions.

face scattering), and even tasks such as imaging and 3D sensing through optically-thin scattering layers.

To demonstrate these capabilities, we build an experimental prototype that we operate outdoors, under direct sunlight and adverse experimental conditions (wind, machine vibrations, vehicle traffic). This is in stark contrast with previous demonstrations of interferometric sensing, which were only possible under carefully-controlled lab conditions (dark room, vibration-isolated optical tables, no air flow). Even under these adverse conditions, our experi-

ments show that it is possible to perform passive depth scanning, at pixel-level lateral resolution and micrometer axial resolution, for objects that are challenging to scan even with active illumination techniques (translucent, metallic, occluded by thin scatterers). More broadly, our results open the door for the deployment of interferometric techniques in uncontrolled outdoor environments, and for the development of passive computational light transport capabilities such as direct-only imaging, imaging through scattering, and transient imaging. To facilitate future research towards these directions, we provide setup details, reconstruction code, and data in the supplement and project website.[1]

Potential impact. The ability to perform interferometry outdoors can be useful for applications such as field inspection (e.g., detecting tiny but dangerous defects on airplane surfaces), field robotics (e.g., high-resolution manipulation), and field medicine (e.g., in battlefields, disaster environments, or impoverished regions). In addition to outdoor operation, all of these applications benefit from *passive* operation: Removing the active source allows for lighter, lower-cost, and lower-power depth sensing, all features critical for field operation—often requiring mobility, small form factors, and functionality under limited electricity access. This is particularly important considering that the source is typically the system component consuming the most power, and also one of the costliest and bulkiest.

2. Related work

Passive depth sensing. A variety of techniques in computer vision measure depth while relying only on external light to illuminate the scene and provide signal to the camera. These techniques use depth cues such as disparity [10, 37, 60], focus and defocus [5, 30, 38, 39, 76], or shading [35, 42]. Unfortunately, all of these techniques are severely constrained in terms of the reflectance properties of the scene.

Active depth sensing. To alleviate these limitations, active depth sensing techniques use a controlled light source to inject light into the scene. Examples techniques include photometric stereo [81, 84], structured light [16, 33, 63, 66, 73], impulse time-of-flight [6, 27, 31, 32, 40, 48, 49, 55, 62, 64, 77, 78], correlation time-of-flight [9, 41, 51, 52, 68, 72, 75], or combinations thereof [1]. However, with a few exceptions [1, 63], active depth sensing techniques fail under strong ambient lighting conditions that overwhelms the projected illumination. Additionally, their increased power consumption due to the light source can make them impractical for power-constrained applications or field operation.

Interferometric depth sensing. Most of the techniques listed above are limited to depth resolutions in the order of millimeters (for passive) or hundreds of micrometers (for active). For applications requiring micrometer depth resolution, one option is to use interferometric techniques [3, 17,

18, 22, 23, 26, 28, 36, 43, 44, 46, 50, 53, 54, 56, 59, 79, 83]. Unfortunately, all these techniques are active, and extremely sensitive to ambient light. Additionally, the micron-scale sensitivity of interferometric techniques makes them notoriously sensitive to environmental conditions such as vibrations and wind, making outdoor operation seemingly impossible. As a consequence, most interferometric techniques are restricted to controlled lab conditions.

Mitigating indirect illumination. Active depth sensing techniques typically operate under an assumption of direct-only illumination. Indirect illumination effects, such as interreflections and subsurface scattering, introduces spurious signal into measurements, resulting in inaccurate depth estimation. To mitigate these effects, *computatioanl light transport* techniques focus on *optically* removing indirect illumination by *probing light transport* [67]. Probing techniques use spatial modulation of the illumination and camera, and include schemes such as epipolar imaging [63, 66], high-spatial-frequency illumination [61, 70], and spatio-temporal code illumination [1, 34, 65]. Similar probing capabilities can be implemented in interferometric systems, by exploiting the spatio-temporal coherence properties of the illumination [28, 50]. Importantly, with the exception of Nayar et al. [61], the ability to suppress indirect illumination is currently only available to active techniques.

Passive interferometry. Passive interferometry has been studied extensively for applications other than outdoor depth sensing. Examples include its use in seismology for subsurface imaging [19, 24, 74], and in ultrasound imaging [80]. More recently, passive interferometry has been used to perform light source localization for occluded imaging applications, in carefully-controlled indoor lab environments [7, 8, 11, 12, 21]. In radio astronomy, passive interferometry techniques use the Hanbury-Brown-Twiss effect to image astronomically distant light sources [14, 15, 45]. Lastly, very-long-baseline interferometry is used for high-resolution imaging of cosmic radio sources, and was used to produce the first-ever image of a black hole [4, 13].

3. Theory of sunlight interferometry

We begin by presenting an analysis of interferometric image formation model using sunlight. Our analysis leverages results from previous works on interferometry with incoherent illumination [28, 50].

The Michelson interferometer. Our optical setup is a version of the classical Michelson interferometer (Figure 2(b)). The interferometer receives illumination from direct sunlight, through an optical configuration that reflects sunlight to be parallel with the optical axis (Section 5). A beam splitter splits the input illumination into two beams: one propagates toward the *scene arm*, and the other toward the *reference arm*. The reference arm is a planar mirror mounted on a translation stage that can vary the mirror's distance from the beam splitter. After reflection, the two light beams re-
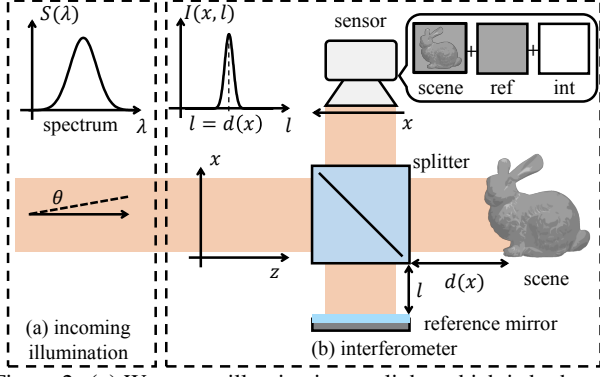
Figure 2. **(a)** We use as illumination sunlight, which is both temporally incoherent (large spectral bandwidth) and spatially incoherent (large angular bandwidth). **(b)** We inject this illumination into a full-field Michelson interferometer, which we use to perform micrometer-scale direct-only time-of-flight imaging.

combine at the beam splitter and propagate toward a camera with a two-dimensional sensor.

To simplify notation, we analyze the setup of Figure 2 in two dimensions—the extension to three dimensions is straightforward. We use an (x, z) coordinate system (Figure 2(b)), where z is the optical axis of the interferometer, coinciding with the illumination and sensor optical axes. We denote by l and $d(x)$ the axial distance from the beamsplitter of the reference mirror and scene point that pixel x images, respectively. We assume that the camera focuses at depth $z = d(x)$. As l is a controllable parameter, we denote it explicitly. We denote by $\boldsymbol{u}_r(x, l)$ and $\boldsymbol{u}_s(x)$ the fields arriving at sensor pixel x after reflection at the reference and scene arms, respectively. Then, the sensor measures,

$$I(x, l) = \underbrace{|\boldsymbol{u}_s(x)|^2 + |\boldsymbol{u}_r(x, l)|^2}_{\equiv \mathcal{I}(x, l)} + 2 \operatorname{Re}\underbrace{\left\{\boldsymbol{u}_s(x)\boldsymbol{u}_r^*(x, l)\right\}}_{\equiv \mathcal{C}(x, l)}. \quad (1)$$

The first two terms in Equation (1) are the intensities the sensor would measure if it were observing each of the two arms separately. We call their sum the *interference-free* image $\mathcal{I}$. The third term, which we call *interference*, is the real part of the complex *correlation* $\mathcal{C}$ between the reflected scene and reference fields. In the rest of this section, we analyze how the correlation changes as a function of the input illumination, reference arm, and scene arm.

Input illumination. Sunlight, and thus the illumination injected in the interferometer, is temporally and spatially incoherent. *Temporal incoherence* means that the illumination is broadband in spectrum, i.e., comprises multiple *independent* waves propagating with different wavelengths. In practice, as we discuss in Section 5, we use a spectral filter in front of the camera to control temporal incoherence. Rather than the wavelength λ, we will characterize the illumination spectrum using the wavenumber $\kappa \equiv 2\pi/\lambda$. Following Goodman [29], we model the temporally-incoherent illumination from filtered sunlight by assuming that its power at different wavenumbers κ follows a Gaus-

sian distribution with mean $\bar{\kappa}$ and standard deviation Δ_κ:

$$\operatorname{Power}(\kappa) \propto \exp\left(-(\kappa - \bar{\kappa})^2/2\Delta_\kappa^2\right). \quad (2)$$

The standard deviation Δ_κ is the *spectral bandwidth* of the illumination: the smaller its value is, the more monochromatic, and thus temporally coherent, the illumination is.

Spatial incoherence means that the illumination comprises multiple planar wavefronts propagating along directions with different angular offsets from the optical axis z. This is accurate for sunlight, as the Sun is a far-field area emitter subtending (when observed from the Earth) a small solid angle. Following Goodman [29], we model the spatially-incoherent illumination from sunlight by assuming that its power at different angular offsets θ is uniform within a small range Δ_θ, and zero otherwise:

$$\operatorname{Power}(\theta) \propto \mathbf{1}_{[-\Delta_\theta/2, \Delta_\theta/2]}(\theta), \quad (3)$$

The range Δ_θ is the *angular bandwidth* of the illumination: the smaller its value is, the more collimated, and thus spatially coherent, the illumination is.

The transmission matrix. To characterize the correlation, we need to consider how the scene arm reflects incident illumination. For this, we use the scene's (continuous) *transmission matrix* $\mathcal{T}(x, x')$: This complex scalar function equals the field generated at point x on the in-focus plane $z = d(x)$, given as input an impulse field at point x' on the same plane [47, 69]. The transmission matrix is the complex, wave-based analogue of the scene's *light transport matrix* [67]. The diagonal elements of the transmission matrix $\mathcal{T}(x, x)$ equal the wave measured by a confocal system that focuses illumination and sensing at the same point x and depth $d(x)$ in the scene. For a system with coaxial illumination (such as the Michelson interferometer), such elements are dominated by direct illumination and are insensitive to most indirect illumination effects (e.g., interreflections, subsurface scattering) [28, 67].

Correlation. We now can characterize the correlation of Equation (1). Given illumination with spectral bandwidth Δ_κ and angular bandwidth Δ_θ, and a scene with transmission matrix $\mathcal{T}$, Gkioulekas et al. [28] prove that,

$$\mathcal{C}(x, l) = \exp\left(i\bar{\kappa}(d(x) - l)\right) G_{\Delta_\kappa}(d(x) - l)$$
$$\int_{x'} \mathcal{T}(x, x') S_{\Delta_\theta}(x - x') \, \mathrm{d}x', \quad (4)$$

where G_{Δ_κ} and S_{Δ_θ} are the *temporal* and *spatial coherence functions*, respectively,

$$G_{\Delta_\kappa}(\tau) \equiv \exp\left(-(\Delta_\kappa\tau)^2/2\right), \quad S_{\Delta_\theta}(\epsilon) \equiv \operatorname{sinc}(2\epsilon\bar{\kappa}\Delta_\theta). \quad (5)$$

We observe that the temporal and spatial coherence functions are shift-invariant kernels, with widths equal to the *temporal* and *spatial* coherence lengths L_G and L_S,

$$L_G \equiv 1/\Delta_\kappa, \quad L_S \equiv 1/\bar{\kappa}\Delta_\theta. \quad (6)$$

The temporal and spatial coherence lengths are inversely proportional to the illumination's spectral and angular band-

widths Δ_κ and Δ_θ, respectively. As the source becomes more monochromatic and collimated, the temporal and spatial coherence lengths become longer, and the temporal and spatial coherence functions become wider, respectively.

Let us consider the case when the incident illumination is sufficiently temporally and spatially incoherent. Then, the spectral and angular bandwidths Δ_κ and Δ_θ are large enough that we can approximate:

$$G_{\Delta_\kappa}(\tau) \approx \delta(\tau), \quad S_{\Delta_\theta}(\epsilon) \approx \delta(\epsilon). \tag{7}$$

Consequently, the correlation in Equation (4) becomes

$$\mathcal{C}(x, l) \approx \mathcal{T}(x, x)\,\delta(d(x) - l). \tag{8}$$

We observe that we measure non-zero correlation only when the position l of the reference mirror matches the scene depth $d(x)$. Then, the correlation equals the wave due to the direct-illumination-only response of the scene. In practice, there will be significant non-zero correlation only when the depth difference between reference and scene is smaller than the temporal coherence length, $|l - d(x)| < L_G$. Likewise, the correlation will include contributions from elements of the transmission matrix close to the diagonal, corresponding to indirect light paths starting at scene points x' whose distance from the imaged point x is smaller than the spatial coherence length, $\|x' - x\| < L_S$. As we show in Section 6, using sunlight corresponds to temporal and spatial coherence lengths in the order of $10\,\mu m$, suggesting that the approximation of Equation (8) is accurate.

4. Direct-only time-of-flight imaging

We now use the results of Section 3 to derive a procedure for performing direct-only time-of-flight imaging using sunlight interferometry. We assume that we want to scan a depth range equal to D. We discretize this depth range to a set of axial locations $l_{m+1} = l_m + L_G/2$, $m = 1, \ldots, M$, where $M = \lceil 2D/L_G \rceil$. The axial separation between consecutive axial locations is due to the fact that, following the discussion at the end of the previous section, the temporal coherence length L_G determines the resolution at which we can measure differences between reference and scene depth.

We use the translation stage of our setup to move the reference mirror to each of the axial locations l_m. At each location, we use the camera to capture an image $I(x, l_m)$—as we use a two-dimensional sensor, we acquire each such image for all locations x in the field of view in a single exposure. We assume that, from each image $I(x, l_m)$, we can isolate the corresponding correlation $\mathcal{C}(x, l_m)$ term in Equation (1). We detail how to do this later in this section.

Then, from Equations (4) and (8), the function

$$\tau(x, l_m) \equiv |\mathcal{C}(x, l_m)|^2, \tag{9}$$

is approximately equal to the *direct-only transient response* of the scene [28, 65]—up to a change of variables $l = c \cdot t$ converting optical pathlength l to time t through the speed of light c. The approximation is due to the non-zero extent

extract interf. Eqs. (11-13)
argmax interf. Eq. (10)

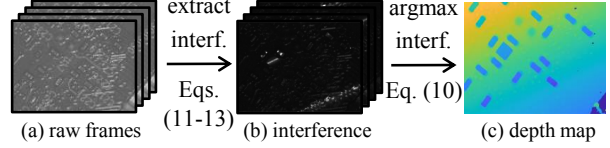(a) raw frames (b) interference (c) depth map

Figure 3. **Processing pipeline.** (a) We capture a stack of images with the reference arm placed at a dense set of positions. (b) We blur the stack temporally and spatially, to estimate interference and correlation amplitude. (c) We extract a depth map by detecting the axial location maximizing correlation amplitude.

of the temporal and spatial coherence functions in Equation (4), proportional to the temporal and spatial coherence lengths. We quantify these coherence lengths, and thus the resolution limits of our technique, in Section 6.

We then estimate depth for each point in the scene as:

$$d(x) = \text{argmax}_{l_m}\,\tau(x, l_m). \tag{10}$$

We also estimate the direct-only intensity as $\tau(x, d(x))$.

We note that Equation (10) assumes that the direct-only transient has a single peak, corresponding to a single reflecting surface at each location x. This assumption is violated, e.g., when the camera observes a semi-transparent object in front of another object. However, in practice, our procedure can handle such cases without issue: As we record the entire transient function $\tau(x, l_m)$, we can search for multiple peaks and detect the depth of all reflecting surfaces. We show this experimentally in Section 7.[2]

Relationship to optical coherence tomography. The process we described above is reminiscent to *time-domain optical coherence tomography (TD-OCT)* [25, 44]. TD-OCT systems use a Michelson interferometer setup, and create incident spatially and temporally-incoherent illumination by placing a source such as an LED or halogen lamp at the focal plane of an illumination lens [28, 82]. Compared to TD-OCT, which requires active illumination, we show that we can perform time-of-flight imaging passively, using sunlight as temporally and spatially incoherent illumination.

Computing correlation. We now describe how to compute the squared correlation amplitude $\tau(x, l_m)$ from the measured axial stack of images $I(x, l_m)$, $m = 1, \ldots, M$. The presence of the complex exponential in Equation (4) suggests that we can use phase shifting [22] to estimate $\tau(x, l_m)$. However, this would require shifting the reference mirror by fractions of the optical wavelength $\bar{\lambda} = 1/\bar{\kappa}$ and capturing 3-4 images around *each* location l_m. This makes acquisition prohibitively long, and is also impractical when operating outdoors—environment vibrations make it difficult to accurately perform sub-wavelength shifts.

Instead, we use the procedure proposed by Gkioulekas

---

[2]We implicitly made the same assumption in Equation (4), when we assumed that the terms depending on l can be moved outside the integral. This corresponds to assuming that the scene's temporally-resolved transmission matrix has the form $\mathcal{T}(x, x', l) = \mathcal{T}(x, x)\,\delta(l - d(x))$. We made this assumption to simplify exposition, but it is not necessary for either our theory or our imaging procedure to work.
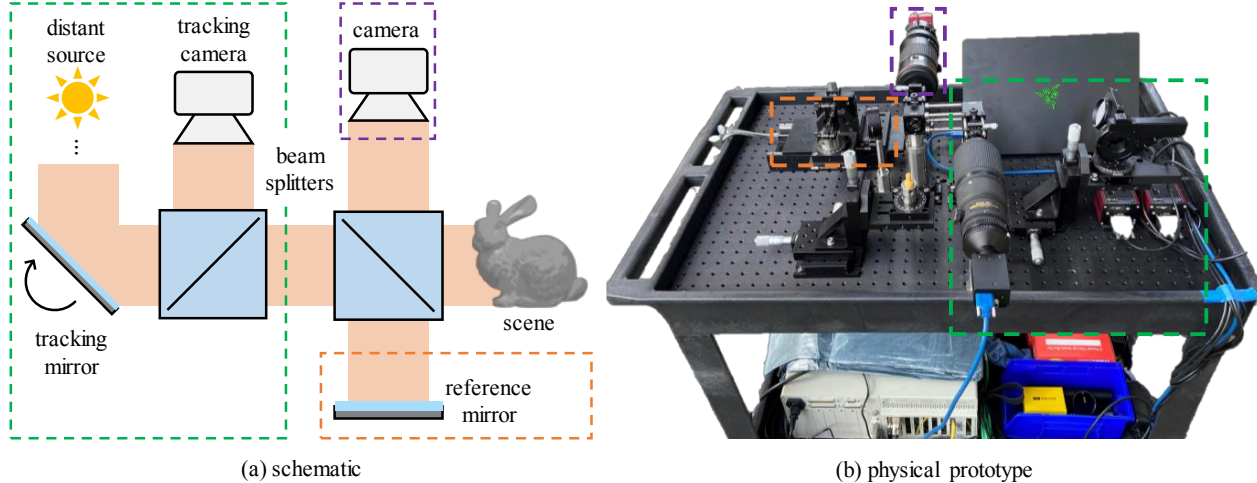
(a) schematic   (b) physical prototype

Figure 4. **(a)** Schematic and **(b)** physical prototype of the sunlight interferometry setup on a utility cart.

et al. [28] to approximate the squared correlation amplitude $\tau(x, l_m)$: First, we estimate the interference-free image $\mathcal{I}(x, l_m)$ by averaging nearby frames in the axial scan,

$$\mathcal{I}(x, l_m) \approx \text{\footnotesize 1/} N \sum_{n=m-N/2}^{N/2} I(x, l_n). \qquad (11)$$

Then, we estimate the squared interference as

$$\text{Re}\{\mathcal{C}(x, l_m)\}^2 \approx \mathcal{R}(x, l_m) = \text{\footnotesize 1/}4(I(x, l_m) - \mathcal{I}(x, l_m))^2. \qquad (12)$$

Lastly, we estimate the squared correlation amplitude as

$$\tau(x, l_m) = (\mathcal{G}_s * \mathcal{R})(x, l_m), \qquad (13)$$

where $\mathcal{G}_s$ is a Gaussian kernel with standard deviation of s pixels. We visualize our pipeline for correlation estimation in Figure 3, and provide pseudocode in the supplement.

The intuition behind the blurring operation in Equation (13) is as follows: When imaging scenes with rough surfaces, the interference appears as speckle. This speckle is created due to pseudo-random sub-wavelength pathlength shifts, accounting for the surface's microstructure. Blurring (squared) interference speckle around some location is approximately equal to averaging squared interference measurements at sub-wavelength shifts. This is how phase shifting techniques estimate squared correlation amplitude [22].

5. Hardware prototype

We build an experimental prototype to demonstrate the capabilities of sunlight interferometry. We show an image of our prototype in Figure 4. We discuss here some key implementation aspects, and refer to the supplement for full implementation details, including a parts list, construction, alignment, and calibration procedures.

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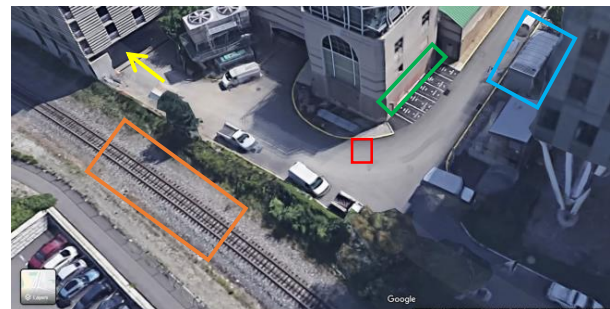

Figure 5. **Google Maps view of experimental site.** This image showcases the challenging conditions of our experiments. **Red**: setup location. **Yellow**: parking garage, causing constant traffic with multiple vehicles (including trucks) passing per acquisition. **Orange**: train tracks, with trains at frequent intervals. **Green**: building air intake causing a constant loud background hum. **Blue**: power transformer also causing a loud constant background hum.
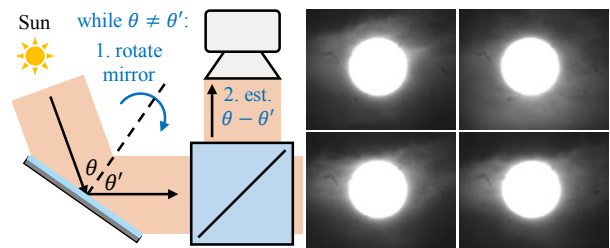
(a) tracking mechanism   (b) tracking camera images

Figure 6. **Sun tracking. (a)** In a closed loop, we estimate the direction of the light beam reflecting from the mirror with a tracking camera focused at infinity, and correct it by rotating the mirror. **(b)** Tracking camera images of the Sun during a centering operation.

resolution $10\,\text{nm}$. We use an axial stepsize of $5\,\mu\text{m}$ (Section 6) and capture 1000 images per scan at exposure times of $50\,\text{ms}$. This results in acquisition times around a minute per scene. We follow Gkioulekas et al. [28] in setting the various imaging settings to maximize interference contrast, and thus signal-to-noise ratio (SNR).

Tracking the Sun. As it is not practical to orient the optical

axis of the interferometer towards the Sun, and to account for motion of the Sun during acquisition, we equip our setup with a custom Sun-tracking assembly (Figure 4, green boxes). This assembly comprises a mirror whose 3D orientation is controlled via two motorized rotation stages. We set the mirror orientation using information from a tracking camera separate from the imaging camera: We set the camera so that its optical axis is parallel to that of the interferometer, and its lens focused at infinity. The camera directly images the Sun through the mirror, and provides feedback (through a digital PID controller) to the rotation stages to adjust the mirror orientation, so that the Sun remains at the center of the camera's field of view.

Outdoor operation. To perform experiments outdoors and mitigate the vibration effects, we build our setup on an optical breadboard mounted on a utility cart. The imaged scenes are on the ground or a tripod. We do not use any enclosures, and both our setup and scenes are fully exposed to ambient light. Figure 5 shows our experimental site. It includes strong sources of environment noise, including moving vehicles, air, and ambient sound. These experimental conditions are in stark contrast with those in prior works on interferometric computational imaging [20, 28, 50, 53, 54, 57], which require carefully-controlled environments (e.g., dark room, vibration-isolated optical tables, no air flow).

Cloudy sky conditions. We performed most of our experiments under direct sunlight. However, we found that under scattered or very high-altitude clouds, we are still able to obtain high-quality depth, even if the clouds are fast-moving. Figure 6(b) shows frames acquired by the tracking camera during a single acquisition, showing the tracked Sun under fast-moving, scattered cloudy conditions.

Cost discussion. Our design trades off the light source of an active system for the Sun tracking module of a passive system. We expect that, in a carefully designed implementation, the passive system will have a significantly lower cost than the active one. The parts we used for tracking in our prototype were simply parts we had readily access to, and can be replaced with much cheaper off-the-shelf components: For example, we can replace the rotation stages with an a pair of galvo mirrors, and the machine vision tracking camera with an inexpensive camera (e.g., Raspberry Pi HQ Camera). Considering the high cost of a light source suitable for OCT (including drivers for current control and temperature stabilization), replacing the source with such a tracking unit reduces overall cost.

Comparison to using an active source. In addition to improved mobility, form factor, power consumption (Section 1) and cost (above), using passive sunlight instead of an active light source for outdoor operation provides more flexible control of imaging performance characteristics. In particular, we can increase axial resolution simply by increasing the bandwidth of the spectral filter we use, with-

out impacting power consumption or signal-to-background ratio—the ratio of sunlight steered through our system to ambient sunlight is unaffected by the filter. By contrast, when using an active light source, it is necessary to also increase the source emission bandwidth to match the increased filter bandwidth. If the source's power remains the same, the result is worse signal-to-background ratio—more ambient light is let through the filter, while active light remains the same. Countering this requires increasing active light, which results in increased power consumption.

6. Coherence properties of filtered sunlight

We measure the temporal and spatial coherence lengths of sunlight, to determine the resolution limits of sunlight interferometry, and inform our experiments in Section 7.

Temporal coherence length. Previous work has estimated the temporal coherence length of direct sunlight to be approximately $L_G = 0.3\,\text{nm}$ [71]. However, performing sunlight interferometry at these axial resolutions using the full spectrum of sunlight is impractical: First, chromatic aberrations from optics (lenses, beamsplitters) when performing interferometry with very broadband illumination can drastically reduce interference contrast [25]. Second, environment vibrations when operating an interferometer outdoors are in the order of hundreds of micrometers. Together, these two factors result in impractically low SNR.

Instead, we limit spectral bandwidth using a spectral filter with central wavelength $550\,\text{nm}$ and bandwidth $20\,\text{nm}$. To measure the resulting temporal coherence length, we perform two experiments where we place either a mirror or a planar diffuser at the scene arm of our setup. For each scene, we measure τ at a dense set of reference axial locations. From Equations (4)-(6), we expect the square root of the measurements to be shaped like a Gaussian with standard deviation equal to the temporal coherence length. We perform a Gaussian fit to the measurements, and estimate the temporal coherence length as the fit's full-width half-maximum. Figure 8 shows the results. For both mirror and planar diffuser, we estimate the temporal coherence length, and thus axial resolution, to be approximately $L_G = 10\,\mu\text{m}$.

Spatial coherence length. To measure the spatial coherence length of sunlight, we use the tracking camera of our setup to capture a direct image of the Sun. At infinity focus, the spatial extent of the Sun in this image is directly proportional to the angular extent of sunlight (up to the focal length of the lens). We can then use Equations (5)-(6) to estimate the spatial coherence length. Figure 9 shows the results. We estimate the angular extent of sunlight to be around $0.57°$, corresponding to a spatial coherence length of approximately $L_S = 100\,\mu\text{m}$. This is consistent with theoretical and experimental estimates in prior work [2, 58].
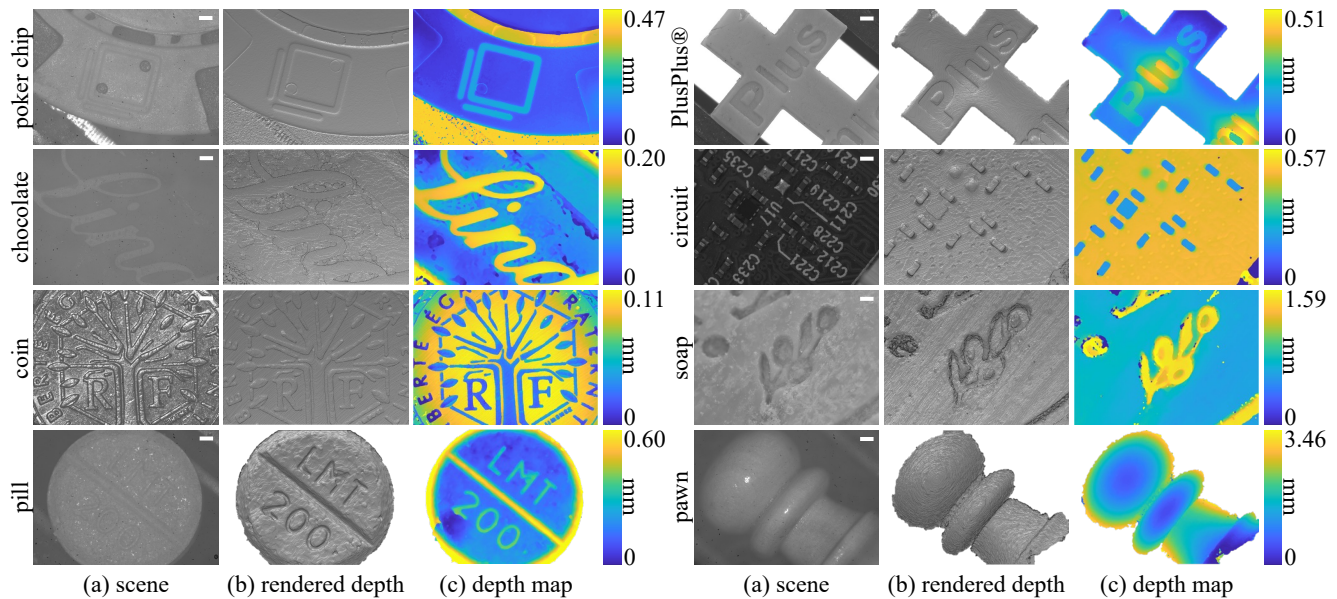
Figure 7. **Depth reconstructions for a variety of challenging scenes using sunlight.** The challenges include environmental factors (Figure 5) and low signal quality (detailed in the text). The white bars on the scene images indicate one millimeter of lateral size.
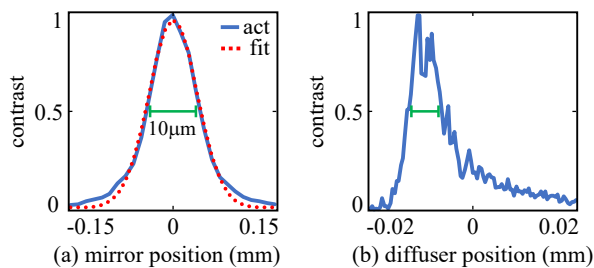

Figure 8. **Temporal coherence of spectrally-filtered sunlight.** We perform Gaussian fits to measurements for both mirror and planar diffuser scenes. From their full-width half-maximum, we estimate temporal coherence length $L_G \approx 10\,\mu m$.
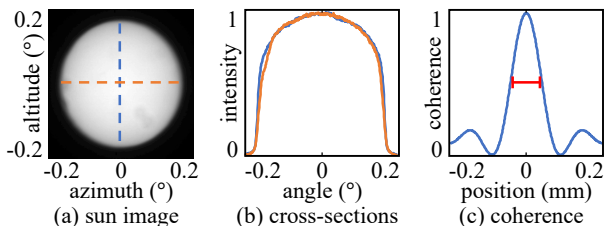

Figure 9. **Spatial coherence length of sunlight. (a-b)** An image of the Sun by the tracking camera yields an angular size of $0.57°$. **(c)** From this, we estimate spatial coherence length $L_S \approx 100\,\mu m$.

7. Experiments

Depth scanning. In Figure 7, we show scans for variety of scenes. These feature micrometer-scale geometric details, and pose a variety of challenges including low reflectivity (circuit board), specularities (coin, pawn) and strong sub-surface scattering (chocolate, soap, pill). We obtain high quality reconstructions, despite operating under very adverse environment conditions (Section 5).

Scanning under extremely low SNR. A key challenge in sunlight interferometry is the very low SNR in captured im-
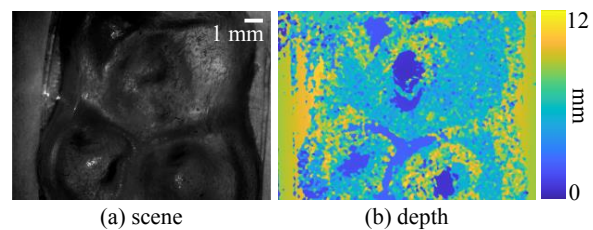

Figure 10. **Depth reconstruction under extremely low SNR.** The gummy bear has very little backreflected direct light, making it extremely challenging to capture depth.

ages. This is for two reasons: First, in Equation (1), the interference is typically much smaller than the interference-free image. Second, the scenes are illuminated not only by the sunlight directed at them by the interferometer, but also by the ambient sunlight, further reducing interference contrast. To test the SNR limits of our technique, in Figure 10, we scan a semi-transparent object (gummy bear) that back-reflects very little light. Despite this, our technique still acquires accurate depth for most of the scene, except for parts observed at near-grazing angles (no backreflected light).

Seeing through diffusers. To test the ability of our technique to isolate direct-only (ballistic) light and measure multiple depth peaks, in Figure 11, we scan a scene where a coin is occluded from the camera by a ground glass diffuser. The conventional images of the scene show strong blur because of the diffuser. Our technique can accurately acquire the depth *and* clean images of the occluded coin.

Time-of-flight with indoor illumination. Even though we focus on sunlight interferometry, Figure 12 shows an experiment testing the ability to perform time-of-flight imaging with passive *indoor* illumination. For this, we point the Sun tracking module to a ceiling light about five feet from the
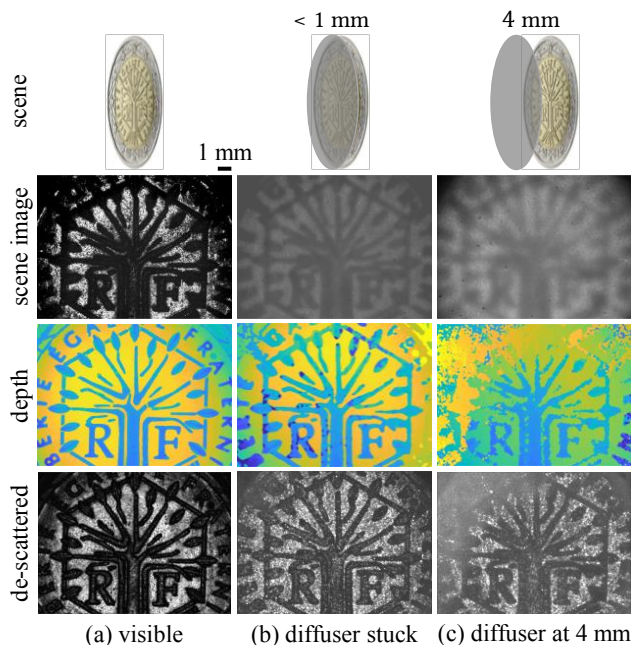
Figure 11. **Seeing through diffusers.** We image a metallic coin under three conditions: **(a)** directly, **(b)** with a ground glass diffuser stuck to it and **(c)** with the diffuser placed 4 mm away. The images of the scene in **(b-c)** are blurred due to scattering in the diffuser. The interference component rejects scattering.
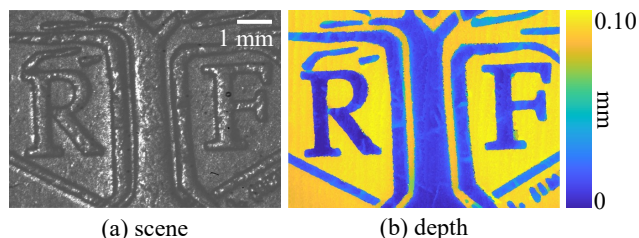

Figure 12. **Depth reconstruction with indoor illumination** at a distance of about five feet from the interferometer.

tracking mirror. We scan the same metallic coin as in Figure 7, which backreflects a lot of light, resulting in accurate depth recovery. However, this scanning setting remains challenging for non-metallic objects, because of the very low SNR between backreflected versus ambient light.

8. Limitations and conclusion

Mechanical scanning and comparison to Fourier-domain and swept-source OCT. As we mentioned before, the operation of our technique is similar to that of (active illumination) time-domain OCT techniques. Both techniques require mechanical scanning along the axial dimension. This results in relatively slow acquisition times (around a minute for our scenes), and increased mechanical vibrations that reduce resolution and SNR. In OCT, these issues are ameliorated through the use of alternative technologies that eliminate the need for axial scanning, such as swept-source and Fourier-domain OCT. Swept-source OCT requires sweeping the wavelength of a monochromatic

source, which is not possible when working with sunlight. Fourier-domain OCT uses a spectrometer to measure interference as a function of wavelength, which relates to the transient response through a Fourier transform. Unfortunately, as two-dimensional spectrometers of sufficiently high spectral resolution are not available, this requires using raster scanning to acquire full-frame images. Additionally, most implementations of Fourier-domain OCT used single-mode fiber optics, which are incompatible with spatially-incoherent sunlight. The development of Fourier-domain sunlight interferometry is an exciting future direction.

General light transport probing. Our method relies on the spatial incoherence of sunlight to capture direct-only measurements of the scene. However, modern computational light transport imaging systems are capable of much more general forms of light transport probing, which currently remain out of reach for sunlight interferometry. Kotwal et al. [50] showed that, by modulating the spatial and temporal coherence properties of the incident illumination, it is possible to implement very general forms of probing inside interferometric imaging systems. Similar forms of modulation can, conceivably, be incorporated in sunlight interferometry. In turn, this would make it possible to realize *passive* variants of computational light transport techniques such as focusing inside scattering and non-line-of-sight imaging. We hope that our work will motivate and facilitate future research towards these fascinating possibilities.

Conclusion. We have demonstrated a passive interferometric method for micron-scale depth sensing outside the controlled environment of an optics lab. Using light from ambient sources such as the Sun removed the need to use expensive, power-hungry, high-power or eye-unsafe lasers for interferometry. The inherent incoherence properties of sunlight made it possible to acquire time-of-flight information at micrometer scales, and insensitive to indirect illumination effects that confound depth inference. Additionally, engineering a robust optical setup mitigated most of the detrimental effects associated with performing experiments in uncontrolled outdoor environments, such as vibrations and wind, that typically hinder interferometric measurements. Together, these contributions allowed us to demonstrate, for the first time, passive imaging capabilities such as micrometer-scanning of challenging scenes, direct-only imaging, and imaging through diffusers. We also take first steps towards enabling these capabilities under passive *indoor* illumination. We hope that our work will motivate applications in areas such as inspection, fabrication, robotic grasping, and biomedical imaging.

References

[1] Supreeth Achar, Joseph R Bartels, William L Whittaker, Kiriakos N Kutulakos, and Srinivasa G Narasimhan. Epipolar time-of-flight imaging. *ACM TOG*, 2017. 2

[2] Girish S. Agarwal, Greg Gbur, and Emil Wolf. Coherence properties of sunlight. *Opt. Lett.*, 29(5):459–461, Mar 2004. 6

[3] Aaron D. Aguirre, Chao Zhou, Hsiang-Chieh Lee, Osman O. Ahsen, and James G. Fujimoto. *Optical Coherence Microscopy*, pages 865–911. Springer International Publishing, Cham, 2015. 2

[4] Kazunori Akiyama, Katherine Bouman, and David Woody. First m87 event horizon telescope results. i. the shadow of the supermassive black hole. *Astrophysical Journal Letters*, 875, 04 2019. 2

[5] Emma Alexander, Qi Guo, Sanjeev Koppal, Steven Gortler, and Todd Zickler. Focal flow: Measuring distance and velocity with defocus and differential motion. In *European conference on computer vision*, pages 667–682. Springer, 2016. 2

[6] Brian Aull. 3d imaging with geiger-mode avalanche photodiodes. *Opt. Photon. News*, 2005. 2

[7] Amaury Badon, Geoffroy Lerosey, Albert C. Boccara, Mathias Fink, and Alexandre Aubry. Retrieving time-dependent green's functions in optics with low-coherence interferometry. *Phys. Rev. Lett.*, 114:023901, Jan 2015. 2

[8] Amaury Badon, Dayan Li, Geoffroy Lerosey, A. Claude Boccara, Mathias Fink, and Alexandre Aubry. Spatio-temporal imaging of light transport in highly scattering media under white light illumination. *Optica*, 3(11):1160–1166, Nov 2016. 2

[9] Seung-Hwan Baek, Noah Walsh, Ilya Chugunov, Zheng Shi, and Felix Heide. Centimeter-wave free-space neural time-of-flight imaging. *ACM Transactions on Graphics (TOG)*, 2022. 2

[10] Stephen T. Barnard and William B. Thompson. Disparity analysis of images. *IEEE TPAMI*, 1980. 2

[11] M. Batarseh, S. Sukhov, Z. Shen, H. Gemar, R. Rezvani, and A. Dogariu. Passive sensing around the corner using spatial coherence. *Nature Communications*, 9(1):3629, Sep 2018. 2

[12] Jeremy Boger-Lombard and Ori Katz. Passive optical time-of-flight for non line-of-sight localization. *Nature Communications*, 10(1):3343, Jul 2019. 2

[13] Katherine L. Bouman, Michael D. Johnson, Daniel Zoran, Vincent L. Fish, Sheperd S. Doeleman, and William T. Freeman. Computational imaging for vlbi image reconstruction. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 913–922, 2016. 2

[14] R. Hanbury Brown and R.Q. Twiss. Lxxiv. a new type of interferometer for use in radio astronomy. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 45(366):663–682, 1954. 2

[15] R. Hanbury Brown and R. Q. Twiss. Correlation between photons in two coherent beams of light. *Nature*, 177(4497):27–29, Jan 1956. 2

[16] Tongbo Chen, Hans-Peter Seidel, and Hendrik P. A. Lensch. Modulated phase-shifting for 3d scanning. In *IEEE/CVF CVPR*, 2008. 2

[17] Yeou-Yen Cheng and James C. Wyant. Two-wavelength phase shifting interferometry. *Applied Optics*, 1984. 2

[18] Yeou-Yen Cheng and James C. Wyant. Multiple-wavelength phase-shifting interferometry. *Applied Optics*, 1985. 2

[19] Jon F. Claerbout. Synthesis of a layered medium from its acoustic transmission response. *GEOPHYSICS*, 33(2):264–269, 1968. 2

[20] Oliver Cossairt, Nathan Matsuda, and Mohit Gupta. Digital refocusing with incoherent holography. *IEEE ICCP*, 2014. 6

[21] Matthieu Davy, Mathias Fink, and Julien de Rosny. Green's function retrieval and passive imaging from correlations of wideband thermal radiations. *Phys. Rev. Lett.*, 110:203901, May 2013. 2

[22] Peter de Groot. *Phase Shifting Interferometry*. 2011. 2, 4, 5

[23] Peter de Groot and John McGarvey. Chirped synthetic-wavelength interferometry. *Optics Letters*, 1992. 2

[24] Deyan Draganov and Elmer Ruigrok. Passive seismic interferometry for subsurface imaging. 2015. 2

[25] Adolf F Fercher, Wolfgang Drexler, Christoph K Hitzenberger, and Theo Lasser. Optical coherence tomography-principles and applications. *Reports on progress in physics*, 2003. 4, 6

[26] Adolf F. Fercher, H Z. Hu, and U. Vry. Rough surface interferometry with a two-wavelength heterodyne speckle interferometer. *Applied Optics*, 1985. 2

[27] Genevieve Gariepy, Nikola Krstajić, Robert Henderson, Chunyong Li, Robert R Thomson, Gerald S Buller, Barmak Heshmat, Ramesh Raskar, Jonathan Leach, and Daniele Faccio. Single-photon sensitive light-in-fight imaging. *Nature Comm.*, 2015. 2

[28] Ioannis Gkioulekas, Anat Levin, Frédo Durand, and Todd Zickler. Micron-scale light transport decomposition using interferometry. *ACM TOG*, 2015. 2, 3, 4, 5, 6

[29] Joseph W Goodman. *Statistical Optics*. Wiley-Interscience, 2000. 3

[30] P. Grossmann. Depth from focus. *PRL*, 1987. 2

[31] Anant Gupta, Atul Ingle, and Mohit Gupta. Asynchronous single-photon 3d imaging. In *ICCV*, 2019. 2

[32] Anant Gupta, Atul Ingle, Andreas Velten, and Mohit Gupta. Photon-flooded single-photon 3d cameras. In *IEEE/CVF CVPR*, 2019. 2

[33] Mohit Gupta, Amit Agrawal, Ashok Veeraraghavan, and Srinivasa G Narasimhan. Structured light 3d scanning in the presence of global illumination. In *IEEE/CVF CVPR*, 2011. 2

[34] Mohit Gupta, Shree K Nayar, Matthias B Hullin, and Jaime Martin. Phasor imaging: A generalization of correlation-based time-of-flight imaging. *ACM TOG*, 2015. 2

[35] Yudeog Han, Joon-Young Lee, and In So Kweon. High quality shape from a single rgb-d image under uncalibrated natural illumination. In *IEEE ICCV*, 2013. 2

[36] Parameswaran Hariharan. *Optical interferometry*. Elsevier, 2003. 2

[37] Richard Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. 2004. 2

[38] Samuel W. Hasinoff and Kiriakos N. Kutulakos. Confocal stereo. In Aleš Leonardis, Horst Bischof, and Axel Pinz, editors, *Computer Vision – ECCV 2006*, pages 620–634, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg. 2

[39] Caner Hazirbas, Sebastian Georg Soyer, Maximilian Christian Staab, Laura Leal-Taixé, and Daniel Cremers. Deep depth from focus, 2018. 2

[40] Felix Heide, Steven Diamond, David B. Lindell, and Gordon Wetzstein. Sub-picosecond photon-efficient 3d imaging using single-photon sensors. *Scientific Reports*, 2018. 2

[41] Felix Heide, Matthias B Hullin, James Gregson, and Wolfgang Heidrich. Low-budget transient imaging using photonic mixer devices. *ACM TOG*, 2013. 2

[42] Berthold Klaus Paul Horn. Shape from shading: A method for obtaining the shape of a smooth opaque object from one view. Technical report, 1970. 2

[43] Kamel Houairi and Frédéric Cassaing. Two-wavelength interferometry: extended range and accurate optical path difference analytical estimator. *J. Opt. Soc. Am.*, 2009. 2

[44] David Huang, Eric A Swanson, Charles P Lin, Joel S Schuman, William G Stinson, Warren Chang, Michael R Hee, Thomas Flotte, Kenton Gregory, Carmen A Puliafito, and James G Fujimoto. Optical coherence tomography. *Science*, 1991. 2, 4

[45] R. C. Jennison. A Phase Sensitive Interferometer Technique for the Measurement of the Fourier Transforms of Spatial Brightness Distributions of Small Angular Extent. *Monthly Notices of the Royal Astronomical Society*, 118(3):276–284, 06 1958. 2

[46] Jon L. Johnson, Timothy D. Dorney, and Daniel M. Mittleman. Enhanced depth resolution in terahertz imaging using phase-shift interferometry. *Applied Physics Letters*, 2001. 2

[47] Benjamin Judkewitz, Roarke Horstmeyer, Ivo M Vellekoop, Ioannis N Papadopoulos, and Changhuei Yang. Translation correlations in anisotropically scattering media. *Nature physics*, 2015. 3

[48] Ahmed Kirmani, Tyler Hutchison, James Davis, and Ramesh Raskar. Looking around the corner using transient imaging. In *IEEE/CVF CVPR*, 2009. 2

[49] Ahmed Kirmani, Dheera Venkatraman, Dongeek Shin, Andrea Colaço, Franco N. C. Wong, Jeffrey H. Shapiro, and Vivek K Goyal. First-photon imaging. *Science*, 2014. 2

[50] Alankar Kotwal, Anat Levin, and Ioannis Gkioulekas. Interferometric transmission probing with coded mutual intensity. *ACM TOG*, 2020. 2, 6, 8

[51] Robert Lange and Peter Seitz. Solid-state time-of-flight range camera. *IEEE JQE*, 2001. 2

[52] Robert Lange, Peter Seitz, Alice Biber, and Stefan Lauxtermann. Demodulation pixels in ccd and cmos technologies. *SPIE*, 2000. 2

[53] Fengqiang Li, Florian Willomitzer, Prasanna Rangarajan, Mohit Gupta, Andreas Velten, and Oliver Cossairt. Sh-tof: Micro resolution time-of-flight imaging with superheterodyne interferometry. *IEEE ICCP*, 2018. 2, 6

[54] Fengqiang Li, Joshua Yablon, Andreas Velten, Mohit Gupta, and Oliver Cossairt. High-depth-resolution range imaging with multiple-wavelength superheterodyne interferometry using 1550-nm lasers. *Applied Optics*, 2017. 2, 6

[55] David B Lindell, Matthew O'Toole, and Gordon Wetzstein. Single-photon 3d imaging with deep sensor fusion. *ACM TOG*, 2018. 2

[56] Sheng-Hua Lu and Cheng-Chung Lee. Measuring large step heights by variable synthetic wavelength interferometry. *Measurement Science and Technology*, 2002. 2

[57] Tomohiro Maeda, Achuta Kadambi, Yoav Y Schechner, and Ramesh Raskar. Dynamic heterodyne interferometry. *IEEE ICCP*, 2018. 6

[58] Heylal Mashaal, Alex Goldstein, Daniel Feuermann, and Jeffrey M. Gordon. First direct measurement of the spatial coherence of sunlight. *Opt. Lett.*, 37(17):3516–3518, Sep 2012. 6

[59] Karl Meiners-Hagen, René Schödel, Florian Pollinger, and Ahmed Abou-Zeid. Multi-wavelength interferometry for length measurements using diode lasers. 2009. 2

[60] Lazaros Nalpantidis, Georgios Sirakoulis, and Antonios Gasteratos. Rev. of stereo vision algorithms: Software to hardware. *Int. J. Optomechatronics*, 2008. 2

[61] Shree Nayar, Gurunandan Krishnan, Michael Grossberg, and Ramesh Raskar. Fast separation of direct and global components of a scene using high frequency illumination. *ACM TOG*, 2006. 2

[62] Cristiano Niclass, Alexis Rochas, Pierre. Besse, and Edoardo Charbon. Design and characterization of a cmos 3-d image sensor based on single photon avalanche diodes. *IEEE JSSC*, 2005. 2

[63] Matthew O'Toole, Supreeth Achar, Srinivasa G Narasimhan, and Kiriakos N Kutulakos. Homogeneous codes for energy-efficient illumination and imaging. *ACM TOG*, 2015. 2

[64] Matthew O'Toole, Felix Heide, David B Lindell, Kai Zang, Steven Diamond, and Gordon Wetzstein. Reconstructing transient images from single-photon sensors. *IEEE/CVF CVPR*, 2017. 2

[65] Matthew O'Toole, Felix Heide, Lei Xiao, Matthias B Hullin, Wolfgang Heidrich, and Kiriakos N Kutulakos. Temporal Frequency Probing for 5D Transient Analysis of Global Light Transport. *ACM TOG*, 2014. 2, 4

[66] Matthew O'Toole, John Mather, and Kiriakos N. Kutulakos. 3d shape and indirect appearance by structured light transport. *IEEE TPAMI*, 2016. 2

[67] Matthew O'Toole, Ramesh Raskar, and Kiriakos N Kutulakos. Primal-dual Coding to Probe Light Transport. *ACM TOG*, 2012. 2, 3

[68] Dario Piatti, Fabio Remondino, and David Stoppa. *State-of-the-Art of TOF Range-Imaging Sensors*. 2013. 2

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

[70] Dikpal Reddy, Ravi Ramamoorthi, and Brian Curless. Frequency-space Decomposition and Acquisition of Light

Transport Under Spatially Varying Illumination. *IEEE ECCV*, 2012. 2

[71] Berke Vow Ricketti, Erik M Gauger, and Alessandro Fedrizzi. The coherence time of sunlight in the context of natural and artificial light-harvesting. *Scientific Reports*, 12(1):1–9, 2022. 6

[72] Alexis Rochas, Michael Gosch, Alexandre Serov, Pierre Besse, Rade S. Popovic, T. Lasser, and Rudolph Rigler. First fully integrated 2-d array of single-photon detectors in standard cmos technology. *IEEE PTL*, 2003. 2

[73] Daniel Scharstein and Richard Szeliski. High-accuracy stereo depth maps using structured light. In *IEEE/CVF CVPR*, 2003. 2

[74] Gerard Thomas Schuster. *Seismic Interferometry*. Cambridge University Press, 2009. 2

[75] Rudolf Schwarte, Zhanping Xu, Horst-Guenther Heinol, Joachim Olk, Ruediger Klein, Bernd Buxbaum, Helmut Fischer, and Juergen Schulte. New electro-optical mixing and correlating sensor: facilities and applications of the photonic mixer device (PMD). In *Sensors, Sensor Systems, and Sensor Data Processing*, 1997. 2

[76] Murali Subbarao and Gopal Surya. Depth from defocus: A spatial domain approach. *IJCV*, 1994. 2

[77] Andreas Velten, Thomas Willwacher, Otkrist Gupta, Ashok Veeraraghavan, Moungi G. Bawendi, and Ramesh Raskar. Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging. *Nature Comm.*, 2012. 2

[78] Federica Villa, Rudi Lussana, Danilo Bronzi, Simone Tisa, Alberto Tosi, Franco Zappa, Alberto Dalla Mora, Davide Contini, Daniel Durini, Sasha Weyers, and Werner Brockherde. Cmos imager with 1024 spads and tdcs for single-photon timing and 3-d time-of-flight. *IEEE JSTQE*, 2014. 2

[79] Zhao Wang, Benjamin Potsaid, Long Chen, Chris Doerr, Hsiang-Chieh Lee, Torben Nielson, Vijaysekhar Jayaraman, Alex E. Cable, Eric Swanson, and James G. Fujimoto. Cubic meter volume optical coherence tomography. *Optica*, 3(12):1496–1503, Dec 2016. 2

[80] Richard L. Weaver and Oleg I. Lobkis. Ultrasonics without a source: Thermal fluctuation correlations at mhz frequencies. *Phys. Rev. Lett.*, 87:134301, Sep 2001. 2

[81] Robert J. Woodham. Photometric Method For Determining Surface Orientation From Multiple Images. *Optical Engineering*, 1980. 2

[82] Peng Xiao, Mathias Fink, and A Claude Boccara. Full-field spatially incoherent illumination interferometry: a spatial resolution almost insensitive to aberrations. *Optics letters*, 2016. 4

[83] Liping Yan, Benyong Chen, Enzheng Zhang, Shihua Zhang, and Ye Yang. Precision measurement of refractive index of air based on laser synthetic wavelength interferometry with edlén equation estimation. *Rev. Scientific Instruments*, 2015. 2

[84] Todd E Zickler, Peter N Belhumeur, and David J Kriegman. Helmholtz stereopsis: Exploiting reciprocity for surface reconstruction. *IJCV*, 2002. 2