Coherent Optical Modems for Full-Wavefield Lidar

PARSA MIRDEHGHAN, University of Toronto, Canada and Vector Institute, Canada
BRANDON BUSCAINO, Ciena Corporation, United States of America
MAXX WU, University of Toronto, Canada and Vector Institute, Canada
DOUG CHARLTON, Ciena Corporation, Canada
MOHAMMAD E. MOUSA-PASANDI, Ciena Corporation, Canada
KIRIAKOS N. KUTULAKOS, University of Toronto, Canada and Vector Institute, Canada
DAVID B. LINDELL, University of Toronto, Canada and Vector Institute, Canada
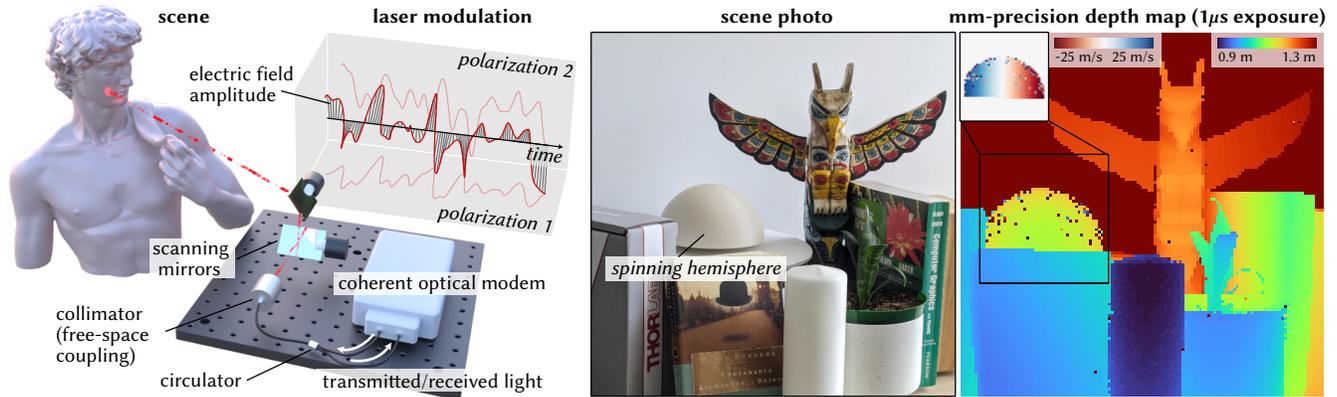
Fig. 1. Overview of full-wavefield lidar. **(left)** Our approach repurposes an off-the-shelf coherent optical modem—typically used for telecommunications—for coherent lidar. The modem modulates the amplitude and phase of light from a 1550 nm laser in two linear polarization states. The light is emitted through a fiber optic cable, free-space collimator, and scanning mirrors, and illuminates a target. The reflected light is coupled into the fiber and directed to the receiver through a circulator. The modem, in this configuration, uses homodyne interferometry to measure the amplitude and phase of light in orthogonal polarization states. **(right)** Based on these measurements, we jointly estimate mm-scale 3D geometry and velocity of dynamic objects with just 1 μs per-pixel exposure time and an eye-safe transmit power of 10 mW. Depth and velocity maps are acquired with 2 mm and 0.9 m/s resolution.

The advent of the digital age has driven the development of coherent optical modems—devices that modulate the amplitude and phase of light in multiple polarization states. These modems transmit data through fiber optic cables that are thousands of kilometers in length at data rates exceeding one terabit per second. This remarkable technology is made possible through near-THz-rate programmable control and sensing of the full optical wavefield. While coherent optical modems form the backbone of telecommunications networks around the world, their extraordinary capabilities also provide unique opportunities for imaging. Here, we repurpose off-the-shelf coherent optical modems to introduce full-wavefield lidar: a type of random modulation continuous wave lidar that simultaneously measures depth, axial

velocity, and polarization. We demonstrate this modality by combining a 74 GHz-bandwidth coherent optical modem with free-space coupling optics and scanning mirrors. We develop a time-resolved image formation model for this system and formulate a maximum-likelihood reconstruction algorithm to recover depth, velocity, and polarization information at each scene point from the modem's raw transmitted and received symbols. Compared to existing lidars, full-wavefield lidar promises improved mm-scale ranging accuracy from brief, microsecond exposure times, reliable velocimetry, and robustness to interference from ambient light or other lidar signals.

CCS Concepts: • **Computing methodologies** → **Computer vision**; **Computational photography**; **3D imaging**.

Additional Key Words and Phrases: coherent lidar, coherent optical modems, 3D imaging, Doppler imaging, polarization imaging

Authors' Contact Information: Parsa Mirdehghan, University of Toronto, Canada and Vector Institute, Canada, parsa@cs.toronto.edu; Brandon Buscaino, Ciena Corporation, United States of America, bbuscain@ciena.com; Maxx Wu, University of Toronto, Canada and Vector Institute, Canada, wumaxx@cs.toronto.edu; Doug Charlton, Ciena Corporation, Canada, dcharlto@ciena.com; Mohammad E. Mousa-Pasandi, Ciena Corporation, Canada, mpasandi@ciena.com; Kiriakos N. Kutulakos, University of Toronto, Canada and Vector Institute, Canada, kyros@cs.toronto.edu; David B. Lindell, University of Toronto, Canada and Vector Institute, Canada, lindell@cs.toronto.edu.

1 Introduction

Coherent optical modems are conventionally used to send digital signals over fiber optic cables by modulating the phase and amplitude of coherent light [Ip et al. 2008; Proakis and Salehi 2008; Roberts et al.

2014]. Driven by the ever-increasing demands for higher networking bandwidths, these modems can now modulate and sample light at staggering rates—up to 100 GHz [Roberts et al. 2017]—across two orthogonal linear polarizations simultaneously [Ip and Kahn 2007]. In effect, modern coherent optical modems achieve *near-THz-rate, programmable control and sensing of the full optical wavefield*, with a reliability that already supports communication over optical fibers spanning thousands of kilometers.

The extreme abilities of these devices to manipulate and sense light within a fiber raise the question: how can we leverage off-the-shelf optical modems to advance the state of the art in free-space imaging? As a first step toward addressing this question, we introduce *full-wavefield lidar (FWL)*. This sensing modality falls broadly within the class of random modulation continuous wave (RMCW) lidar; however, compared to conventional RMCW techniques [Sambridge et al. 2021] that use binary phase [Bashkansky et al. 2004] or amplitude [Ai et al. 2011] modulation, our approach employs controllable modulation of the full coherent wavefield across two polarization states. Further, our FWL reconstruction algorithm provides simultaneous estimates of depth and velocity without requiring specialized encoding schemes or trading off bandwidth [Banzhaf and Waldschmidt 2021].

To realize FWL, we use free-space coupling optics and a conventional galvanometer to turn a 400 Gb/s off-the-shelf coherent optical modem[1] into a coherent lidar system that raster-scans the field of view (see Figure 1). We develop a time-resolved image formation model that captures the unique properties of the raw output of optical modems repurposed for free-space imaging—including internal reflections, Doppler shifts, and the scrambled polarization state of back-scattered light—and use this model to formulate a maximum-likelihood reconstruction algorithm. Our algorithm relies on the modem's raw output to solve an inverse problem that jointly recovers depth, velocity, and polarization information.

Compared to existing lidars, FWL promises significantly more flexibility and control over the transmitted waveforms of light; mm-scale ranging; reliable velocimetry; improved performance at very short (e.g., microsecond) exposure times with eye-safe transmit power; and robustness to interference from ambient light or other lidar signals. We demonstrate full-wavefield lidar by capturing a variety of challenging scenes with moving objects, partial transparencies, strong ambient light, and specular surfaces.

2 Background and Related Work

Our work relates to multiple types of lidar and to other sensing modalities. We provide an overview of the connections to incoherent lidar, coherent lidar, and to optical telecommunications technologies.

2.1 Incoherent Lidar

Most commercial lidar systems operate on a principle of incoherent detection. These lidars modulate the intensity of light to recover scene geometry by measuring phase delays of a sinusoidally modulated signal [Coddington et al. 2009; Schuhler et al. 2006] or propagation delays of emitted and backscattered laser pulses [Lin and Liu 2004; Rapp et al. 2020; Takeuchi et al. 1983]. Incoherent lidars can

---

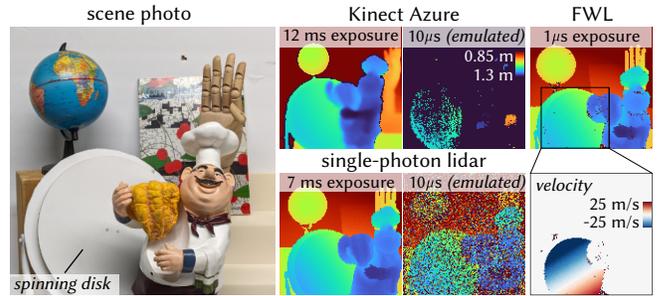[1]Ciena Corporation WaveLogic 5 Nano (https://www.ciena.com/products/wavelogic/wavelogic-5/nano).

Fig. 2. Qualitative comparison of FWL to the Kinect Azure [Bamji et al. 2018] and single-photon lidar. FWL recovers accurate depth and velocity with only 1 μs exposure times per pixel and an eye-safe 2 mW laser. The Kinect fails at light levels corresponding to 10 μs exposure times, which we emulate using neutral density filters. For the single-photon lidar system (see the supplement for a description), we emulate a 10 μs exposure time by thinning the detected photon counts [Lewis and Shedler 1979].

also capture polarization information of backscattered light [Baek and Heide 2021, 2022]. However, incoherent detection schemes are sensitive to interference from other light sources or ambient light (e.g., from other lidars or the sun). While incoherent continuous wave systems can measure velocity from phase shifts due to the Doppler effect [Heide et al. 2015; Hu et al. 2022], pulsed systems are not sensitive to phase information and cannot be used for velocimetry in the same fashion. FWL recovers accurate depth with 1 μs exposure times that are 10,000× shorter than that of incoherent, intensity-modulated depth sensors (e.g., the Kinect Azure [Bamji et al. 2018], which uses ≈10 ms exposure times—see Figure 2). Also, similar to pulsed systems [O'Toole et al. 2017] and some continuous-wave systems [Kadambi et al. 2013; Peters et al. 2015], our approach recovers time-resolved profiles of backscattered light.

Single-photon lidar. Incoherent lidars based on single-photon detection are notable for their extreme sensitivity to individual particles of light [Kirmani et al. 2014; Lindell et al. 2018]. However, the advantages of single-photon lidar are primarily in the weak signal regime where photons arrive infrequently at rates that are much lower than the detector dead time, which is typically on the order of tens of nanoseconds. At higher signal levels, photon arrivals are missed, leading to difficult-to-model, non-linear distortions in photon arrival times that skew ranging estimates [Rapp et al. 2021].

Fig. 3. Illustration of coherent and incoherent detection schemes.

Our system functions robustly with exposures of 1 μs or less; at such short exposure times, a typical single photon lidar in the linear, low-flux regime might detect less than one laser photon on average. Further, any received photons could be obscured by detections from ambient light. This makes single-photon lidar very challenging when dealing with very short exposure times and ambient light.

2.2 Coherent Lidar

Coherent lidar detects the amplitude and/or phase of backscattered incident laser light by interfering it with unmodulated light from
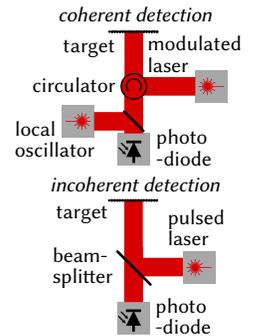
the same laser (referred to as the local oscillator [Henderson et al. 2005]), or from another laser at a different frequency [Li et al. 2021]. In contrast to incoherent lidar or other techniques such as optical coherence tomography [Gkioulekas et al. 2015; Kotwal et al. 2023], it is critical that the laser source have a high degree of temporal coherence so that the incident light and local oscillator remain correlated when they are interfered at a photodiode, as illustrated in Figure 3. Please refer to Goodman [2015] for a detailed treatment of temporal coherence; we provide a mathematical description of coherent detection in the supplement.

In general, coherent lidar systems have several advantages compared to their incoherent counterparts. Since they use continuous wave emission, they can allow eye-safe operation at higher average optical powers compared to pulsed lasers, which may have very high peak power depending on the duty cycle. Additionally, their use of coherent averaging (i.e., of the complex electric field) results in a linear increase in signal-to-noise ratio (SNR) with exposure time compared to the square-root relation of incoherent averaging of intensity measurements [Baumann et al. 2019]. Further, coherent detection strongly suppresses interference from ambient light due to the interferometric detection procedure and use of balanced photodetectors [Behroozpour et al. 2017].

FMCW lidar. Perhaps the most common type of coherent lidar is based on a frequency-modulated continuous wave (FMCW) transmit signal [Zheng 2005]. Specifically, FMCW lidars transmit a "chirp" signal whose optical frequency increases linearly with time. While one advantage of FMCW lidar is the simplicity of its modulation scheme (i.e., the linear frequency sweep), the system performance is strongly coupled to the modulation waveform. FMCW depth resolution is tied to the bandwidth of the linear frequency sweep, and the maximum depth and velocity resolution depend on the duration of the sweep [Behroozpour et al. 2017]. For many swept-source lasers, there is limited flexibility to adjust these parameters (e.g., beyond certain preset configurations). In contrast, our system's performance is easily configurable and largely decoupled from the modulation waveform: the depth resolution is tied to the sample rate (74 GHz for our system), and the maximum depth and the velocity resolution depend on the exposure time—each of which can be continuously controlled. Our setup is also more general as it measures the response of a scene to arbitrary, programmable wavefields. Additionally, FMCW lidars are sensitive to interference from other frequency-modulated lidars [Hwang et al. 2020]. RMCW techniques, including our proposed FWL system, are less susceptible to interference than FMCW because the transmit waveforms are likely to be orthogonal to any interfering signal [Chen et al. 2023; Hwang and Lee 2020].

RMCW lidar. FWL can be considered as a type of RMCW lidar [Takeuchi et al. 1983]; however, different from conventional RMCW lidar, we do not use a binary modulation of phase [Bashkansky et al. 2004; Sambridge et al. 2021] or amplitude [Ai et al. 2011; Takeuchi et al. 1986]. Instead, we use discrete, pseudorandom modulation of both amplitude and phase in two polarization states. In this sense, our work is close to "true random" or "chaos" lidars that use amplified spontaneous emission noise [Hwang and Lee 2020] or chaotic microcombs [Chen et al. 2023] to induce random fluctuations in the amplitude of emitted light. However, unlike

FWL, their output waveforms cannot be programmably controlled since they are governed by stochastic phenomena.

Many of the drawbacks associated with FMCW lidar (e.g., related to configurability or interference as discussed above) can be mitigated with RMCW lidar—though there are considerable implementation challenges. The range resolution is determined, in part, by the modulation speed and sample rate—typically tens of GHz to achieve mm-scale resolution. Thus, few examples of this modality exist in the literature due to the significant hardware requirements related to ultrafast sample rates and the computational challenge of modeling Doppler shifts, laser phase noise, speckle, and polarization changes induced by scattering.

Our work overcomes hardware challenges associated with RMCW lidar by leveraging existing, off-the-shelf optical modems used for telecommunications. We demonstrate that FWL improves the accuracy of ranging and velocimetry over other modulation schemes (e.g., without phase modulation or amplitude modulation) implemented on the same optical modem. We also show that our polarization-aware reconstruction framework improves accuracy and performance at low SNRs compared to matched-filtering schemes similar to those used in RMCW lidar.

2.3 Optical Telecommunications

Our work makes use of coherent optical modems that are conventionally used to send digital signals over fiber optic cables [Ip et al. 2008; Roberts et al. 2014]. Typically, these modems use a modulation scheme [Proakis and Salehi 2008] to optically encode digital information in the amplitude, phase, polarization, and frequency of transmitted light. Signals from optical modems are often combined with wavelength division multiplexing [Brackett 1990] and transmitted in parallel across a single-mode fiber at different wavelengths (e.g., 1530–1565 nm). Polarization division multiplexing is also used to transmit two signals in parallel using orthogonal linear polarizations of the electric field [Ip and Kahn 2007].

We operate a coherent optical modem on a single wavelength channel at 1550 nm; this wavelength has the benefit of being eye safe at roughly 50× higher transmit powers (up to 10 mW) compared to visible wavelengths because light is absorbed at the cornea rather than propagating to the retina [Sliney and Mellerio 2013].

We use a simple modulation scheme consisting of random amplitude and phase values sampled from a complex Gaussian distribution—the optimal scheme for measurements corrupted by Gaussian noise [Forney 1992]. Modulated light is transmitted on two orthogonal polarization channels which are coupled to free space and backscattered from surfaces with different material properties. As such, the transmitted light signals are distorted by speckle, and polarization and phase information is scrambled. We explicitly estimate these distortions to recover depth and velocity.

3 Coherent Optical Modem Imaging

Coherent modems are designed to enable high-bandwidth transmission of data over optical fiber. Achieving data rates of many gigabits per second necessitates exploiting multiple degrees of freedom in the transmitted light. In this section, we outline the working principle of optical modems—from encoding digital information into discrete symbols, to transmitting, receiving, and demodulating the digital
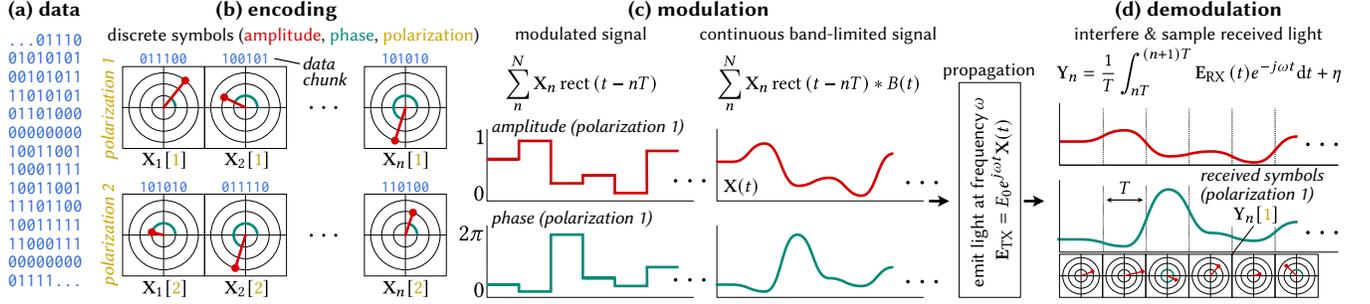
Fig. 4. Overview of data transmission in coherent optical modems. **(a,b)** Binary data are collected and encoded into a discrete sequence of symbols X_n, each paired with a certain amplitude, phase, and polarization state of light. **(c)** This sequence of symbols is used to create a piecewise constant waveform with segments of duration T (shown for a single polarization). In practice, a band-limited version of this waveform modulates laser light with amplitude E_0 and wavelength $\lambda = c\frac{2\pi}{\omega}$, where $\omega/2\pi$ is the optical frequency and c is the speed of light. **(d)** The modulated light is emitted, collected by a receiver, and interfered with an unmodulated receiver-side laser to remove the optical frequency shift. The resulting waveform is sampled to recover the demodulated symbols Y_n.

data. Then, we outline how optical modems can be repurposed for imaging in free space.

3.1 Coherent Modulation and Demodulation

A coherent optical modem realizes two main functionalities: *modulation* of a coherent laser, where a predefined data sequence is encoded into a laser's electric field; and *demodulation* of received light, where the transmitted data sequence is inferred from a measured electric field (see Figure 4).

Modulation. We consider a coherent modem that modulates the phase and amplitude of light in two orthogonal linear polarization states. Given an input sequence of digital data (Figure 4), we can define a complex-valued matrix $X \in \mathbb{C}^{2 \times N}$ that is used to map digital data to different output states of the coherent laser. More specifically, each column of this matrix, denoted by X_n, contains two complex-valued *symbols* that specify the phase and amplitude of the outgoing electric field at each polarization during the nth output time interval, or *symbol period*: $nT \le t < (n+1)T$.

Once the digital data are encoded into symbols, the modulation process involves two steps. First, the discrete sequence of symbols X_n is transformed into a continuous coded waveform $X(t)$:

$$X(t) = \sum_{n=0}^{N-1} X_n \, \text{rect}\left(t - nT\right) * B(t), \qquad (1)$$

where $\text{rect}(t)$ is a rectangular function that is equal to one for $0 \le t < T$ and zero elsewhere, and $B(t)$ is a filter that creates a smooth, band-limited signal from the piecewise concatenation of rectangular functions. Second, the coded waveform modulates a laser of wavelength $\lambda = c\frac{2\pi}{\omega}$ whose electric field propagates at the speed of light c and oscillates with carrier frequency $\omega/2\pi$ and amplitude E_0. The transmitter output of the coherent optical modem can then be formulated as

$$E_{\text{TX}}(t) = E_0 e^{j\omega t} X(t), \qquad (2)$$

such that the digital information is completely encoded in the transmitted electric field E_{TX}.

Demodulation. The goal of demodulation is to recover an estimate Y_n of the transmitted symbol sequence X_n by measuring the amplitude and phase of received electric field $E_{\text{RX}}(t)$. Either *homodyne*

or *heterodyne* coherent detection can be used, where the measured electric field is interfered with a laser source called the *local oscillator*. In the context of optical communications, the transmitted signal and the local oscillator are generated by different laser sources; the phase of the local oscillator is matched to that of the received signal using a phase-locked loop or feed-forward carrier synchronization, which maintains temporal coherence [Ip et al. 2008].

In a homodyne detection scheme, the received signal and local oscillator have the same carrier frequency $\omega/2\pi$. Interference of these two sources *downconverts* the received signal: the high-frequency oscillations of the electric field at the laser frequency are removed, and the modulated waveform containing the encoded digital data is recovered (refer to the supplement for a mathematical description of this procedure). The detected sequence Y_n is given as:

$$Y_n = \frac{1}{T} \int_{nT}^{(n+1)T} E_{\text{RX}}(t) \, e^{-j\omega t} dt + \eta. \qquad (3)$$

Here, the received electric field after interference with the local oscillator is sampled via integration over the symbol period. The term η is complex Gaussian noise that incorporates multiple effects, including thermal noise in the receiver, shot noise, and noise due to inline optical amplifiers [Ip et al. 2008].

We note that this section provides a simplified description of the coherent modulation and demodulation procedure. In practice, optical modems contain additional optical and digital processing stages to ensure that the local oscillator is locked to the transmit laser and that the received signal is sampled with the correct timing. While shot noise generally follows a Poisson distribution, here it is approximated by a Gaussian distribution. Further, we note that the Gaussian noise model neglects secondary effects such as optical fiber non-linearities [Singh and Singh 2007]. For a detailed treatment of coherent detection in optical modems, please refer to Ip et al. [2008].

3.2 Repurposing Coherent Optical Modems for Imaging

Our goal is to repurpose coherent optical modems for the task of free-space 3D imaging. Given a known transmit symbol sequence and the received symbol sequence, we aim to infer unknown scene parameters, namely depth, radial velocity, and change in polarization state. In contrast to communication systems where the transmitted sequence is unknown and the goal is to recover the digital data, we
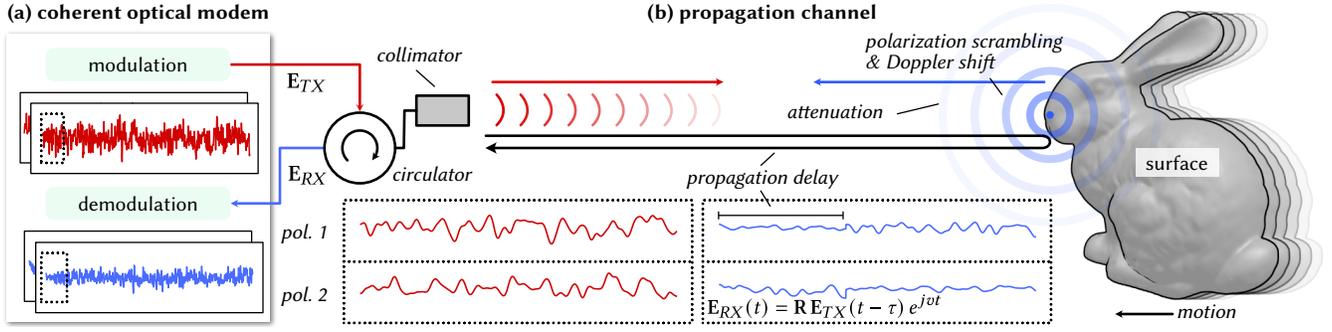
Fig. 5. Illustration of imaging with coherent optical modems. **(a)** The modulated wavefield E_{TX} is transmitted to the target through a fiber optic cable, circulator, and collimator. The received wavefield E_{RX} is demodulated and detected by the optical modem. **(b)** The transmitted wavefield is distorted by multiple effects: the propagation delay induces a shift in the measurements, shown in the bottom plots of E_{RX}; scattering off of a moving surface scrambles the two transmitted polarization channels (modeled by multiplication with a Jones matrix R) and induces a Doppler shift of frequency v; the wavefield is attenuated as it propagates back to the collimator; last, the measurements are corrupted by noise η from the optical modem or optical amplifiers (not shown).

seek to understand how transmitted waveforms are distorted by the propagation channel and, in turn, recover scene properties.

System overview. Figure 5 provides an overview of the proposed system. The optical modem modulates the laser electric field based on an arbitrary, known symbol sequence, and the light is transmitted to the scene through a fiber optic cable and collimator. The reflected light is collected by the same collimator and directed to the optical modem. In our setup, the transmitter and receiver share the same light propagation path through a circulator (which separates optical signals traveling in opposite directions), and we use a set of scanning mirrors to image the scene (Figure 1).

Measurement model. The received and demodulated electric field $E_{RX}(t)$ is altered by three main effects resulting from propagation to the scene and back. First, the demodulated electric field is time-shifted relative to the transmit signal due to the propagation delay to the target and back. Second, if illuminating a moving target, the field is frequency-shifted due to the Doppler effect [McManamon 2015]. Finally, the field is distorted by attenuation and changes in the polarization state due to the surface reflection. The received electric field can therefore be modeled as

$$E_{RX}(t) = R\, E_{TX}(t - \tau)\, e^{jvt}. \tag{4}$$

The terms τ and v denote the propagation delay and frequency shifts of the demodulated electric field. The Jones matrix R is a 2×2 complex-valued matrix that describes how the transmitted electric field is transformed by a polarization-dependent attenuation and rotation induced by the properties of the target surface and propagation through the fiber [Gordon and Kogelnik 2000]). We will later show that the optical modem's measurement of two polarization channels enables recovery of the Jones matrix.

In practice, we assume that the received, downconverted electric field is approximately constant over the symbol integration period, allowing us to drop the integral in Equation 3 and relate the received symbol sequence Y_n directly to the transmitted sequence as

$$Y_n = R\, X_{n-\Delta}\, e^{jvt} + \eta, \tag{5}$$

where $\Delta = \lfloor \tau/T \rfloor$ is the integer shift in the symbol sequence due to the propagation delay. This integer approximation to the time delay (i.e., $\tau/T \approx \lfloor \tau/T \rfloor$) is justified in our case, where $\tau \gg T$.
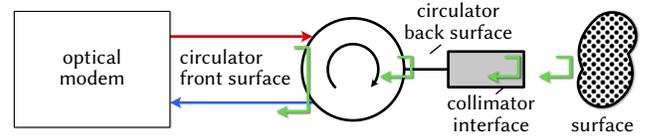

Fig. 6. Our FWL prototype and coaxial lidar systems in general have non-idealities such as reflections from interfaces in the optical path.

Coaxial lidar configurations like ours exhibit inter-reflections caused by the various interfaces in the optical propagation path. Additional reflections from partially transparent surfaces in the scene are also possible. Thus measurements from our system (Figure 5) can be modeled as a superposition of signals, including multiple inter-reflections and scene reflections as depicted in Figure 6. While we are mainly interested in the reflection off of target surfaces, other reflections are captured, e.g., from the circulator interfaces, and from the glass–air interface of the collimator. Since these reflections due to internal surfaces are static, they do not induce a Doppler shift. We therefore consider a generalization of Equation 5 that models the received symbol sequence as a discrete superposition of the delayed, polarization-scrambled, and potentially frequency-shifted copies of the transmitted signal:

$$Y_n = \sum_{s=0}^{S-1} R_s\, X_{n-\Delta_s}\, e^{jv_s t} + \eta. \tag{6}$$

Here, s denotes the index of each copy of the transmitted signal; R_s and $\Delta_s = \lfloor \tau_s/T \rfloor$ model the corresponding polarization scrambling, signal attenuation, and propagation delay; and v_s is the corresponding Doppler shift (zero for internal reflections).

Key objective. Given a known transmitted symbol sequence, our goal is to recover per-pixel scene unknowns from the received symbol sequence through maximum likelihood estimation:

$$\{R_s^*, \Delta_s^*, v_s^*\} = \underset{\{R_s, \Delta_s, v_s\}}{\mathrm{argmax}}\ \Pr(Y_n | X_n). \tag{7}$$

For each pixel, Δ_s^* is an estimate of the depth $d_s = \Delta_s^* \cdot Tc/2$, where c is the speed of light. The velocity is calculated from the Doppler shift as $\frac{v_s^* \cdot c}{\omega}$, and R describes the attenuation and polarization change, which depend on the surface and material.

4 Joint Estimation of Depth, Velocity, and Polarization

We provide an optimization algorithm for joint estimation of depth, radial velocity, and polarization change at each pixel.

Matched filtering. In the case of a single direct reflection ($S = 1$ in Equation 6), no polarization scrambling, and additive Gaussian noise, the matched filter is the maximum-likelihood estimator for recovering the unknown propagation delay of a known transmit waveform [Turin 1960]. The matched filter is also the typical approach for depth estimation in RMCW lidar [Sambridge et al. 2021; Takeuchi et al. 1983]. Using the notation of Equation 7, matched filtering can be expressed as the optimization

$$\Delta^* = \underset{\Delta}{\operatorname{argmax}} \left| \sum_{\Delta=-\infty}^{\infty} \overline{\mathbf{X}}_{n-\Delta} \mathbf{Y}_n \right|^2, \qquad (8)$$

where $\overline{\mathbf{X}}_n$ is the complex conjugate of the transmitted symbol $\mathbf{X}_n$.

Joint estimation. While matched filtering offers a straightforward and computationally efficient solution to depth estimation, it is not well-suited for complex scenes. Specifically, in the general case captured by Equation 6, no single time delay can explain the received symbol sequence because of polarization scrambling, Doppler shift, and multiple reflections. Crucially, by ignoring these effects, the optimization in Equation 8 provides no information about the velocity and polarization properties of scene points. Proper handling of these effects is essential, not only for FWL, but for any lidar system—multiple reflections can be caused by depth discontinuities or partially transparent surfaces. In coherent lidar, Doppler shifts must be modeled for accurate depth estimation and velocimetry for dynamic scenes.

Instead, we seek the values of $\mathbf{R}$, Δ, and ν, as well as the number of shifted copies of the transmitted symbols S, that minimize the mean squared error between the transmitted and received symbols, described as follows:

$$\underset{\{S, \mathbf{R}_s, \Delta_s, \nu_s\}}{\operatorname{argmin}} \sum_{n=0}^{N-1} \left\| \mathbf{Y}_n - \sum_{s=0}^{S-1} \mathbf{R}_s \, \mathbf{X}_{n-\Delta_s} \, e^{j\nu_s t} \right\|_2^2. \qquad (9)$$

Solving this problem provides the maximum likelihood estimate of these parameters under a Gaussian noise model. This approach is analogous to channel estimation employed in the digital communications literature [Proakis and Salehi 2008].

The optimization in Equation 9 does not have a closed-form solution; the associated objective function is combinatorial in nature because of the (typically small and unknown) number of additive terms. To make optimization tractable, we relax the objective by discretizing the space of Doppler shifts and associating an unknown Jones matrix $\mathbf{R}_{\Delta, \nu}$ to each possible time delay Δ and Doppler shift ν.

$$\mathcal{L}_{\text{data}} = \sum_{n=0}^{N-1} \left\| \mathbf{Y}_n - \sum_{\Delta, \nu} \mathbf{R}_{\Delta, \nu} \, \mathbf{X}_{n-\Delta} \, e^{j\nu t} \right\|_2^2. \qquad (10)$$

Since we expect a relatively small number of contributions to the sum of Equation 9, we introduce a sparsity-promoting regularization term $\mathcal{L}_{\text{sparse}}$ into our objective function. We also introduce a total variation spatial regularization term $\mathcal{L}_{\text{TV}}$ to help mitigate errors
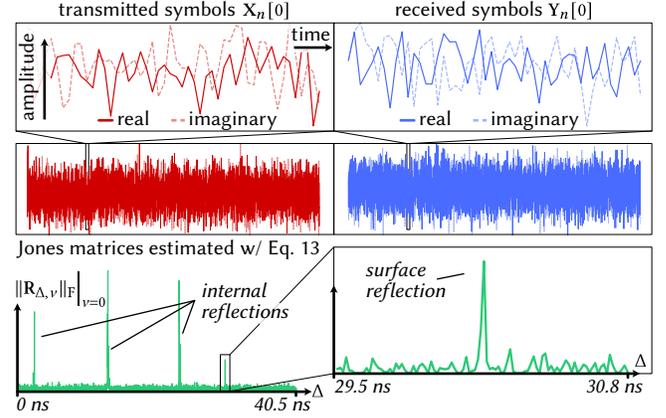
Fig. 7. Estimated Jones matrices. The transmitted and received symbols (top and middle) are related by a convolution with a sequence of Jones matrices $\mathbf{R}_{\Delta, \nu}$ (bottom) estimated using Equation 13. The real and imaginary symbol components (top) describe the amplitude and phase (Fig. 4). We plot $\|\mathbf{R}_{\Delta, \nu}\|_F$ (bottom) for a static scene with $\nu = 0$.

across pixels due to speckle noise [Goodman 2007] by encouraging spatial smoothness in the energy of the reconstructed Jones matrices:

$$\mathcal{L}_{\text{sparse}} = \sum_{\Delta, \nu} \|\mathbf{R}_{\Delta, \nu}\|_F, \text{ and} \qquad (11)$$

$$\mathcal{L}_{\text{TV}} = \sum_{i, j, \Delta, \nu} \sqrt{|(\mathbf{D}_v \|\mathbf{R}_{\Delta, \nu}\|_F)_{i,j}|^2 + |(\mathbf{D}_h \|\mathbf{R}_{\Delta, \nu}\|_F)_{i,j}|^2}. \qquad (12)$$

Here, $\|\cdot\|_F$ is the Frobenius norm, $\mathbf{D}_v$ and $\mathbf{D}_h$ compute vertical and horizontal finite differences between pixels, and i and j index the vertical and horizontal pixel locations. The resulting optimization problem is

$$\underset{\mathbf{R}_{\Delta, \nu}}{\operatorname{argmin}} \, \mathcal{L}_{\text{data}} + \lambda_{\text{sparse}} \mathcal{L}_{\text{sparse}} + \lambda_{\text{TV}} \mathcal{L}_{\text{TV}}, \qquad (13)$$

where the scalars λ_{sparse} and λ_{TV} weigh each regularizer.

Once the Jones matrices have been estimated for all Δ and ν, we solve for depth and velocity as follows. Given that the scene contains a single reflection from a target surface, we return the delay Δ and frequency shift ν whose associated Jones matrix has the maximum Frobenius norm (and we ignore delays due to internal reflections). That is, we find

$$(\Delta^*, \nu^*) = \underset{\Delta, \nu}{\operatorname{argmax}} \|\mathbf{R}_{\Delta, \nu}\|_F \quad \text{subject to} \quad \Delta > \Delta_{\min}, \qquad (14)$$

where $\Delta_{\min}$ is the minimum delay due to free-space propagation, ignoring internal reflections in the optical modem.[2]

4.1 Implementation Details

Optimization. We implement the optimization in Equation 13 in PyTorch [Paszke et al. 2019] and use Adam [Kingma and Ba 2015] with a learning rate of 1×10^{-2}. For the weighting of loss terms we use $\lambda_{\text{sparse}} = 0.1$ for static scenes and $\lambda_{\text{sparse}} = 0.3$ for dynamic scenes, with $\lambda_{\text{TV}} = 0.1$. In practice, to ease computational requirements, we only apply the TV loss to Jones matrices across the Δ dimension where $\nu = 0$ (i.e., to the static scene components).

---

[2] We assume that the values Δ associated with internal reflections are calibrated a priori.
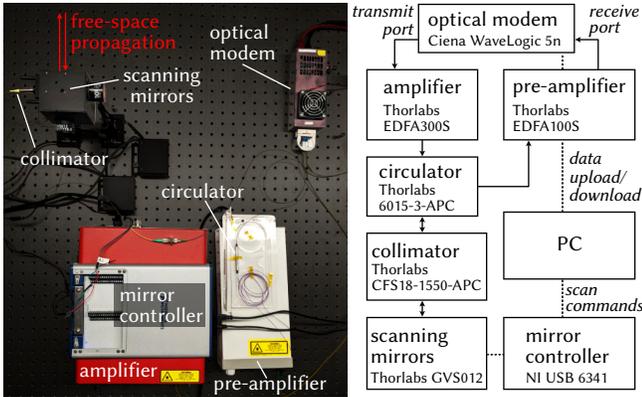
Fig. 8. Prototype coherent lidar system. **(left)** Each component is shown in a photo of the system. **(right)** The connections between each component are illustrated. The transmit and receive ports of the optical modem are coupled into single-mode optical fiber. The transmit cable connects through an amplifier and collimator to a set of scanning mirrors, whose positions are set by a controller. The collimator couples the reflected light through the circulator, to the pre-amplifier, and into the receive port of the optical modem. A PC interfaces with the modem and the scanning mirrors.

Similarly, for dynamic scenes we set a maximum axial velocity of 30 meters/second, and only optimize for the feasible frequency shifts. Please see the supplement for additional implementation details related to the optimization.

Hardware prototype. A photo and illustration of the hardware prototype are shown in Figure 8. We use a Ciena WaveLogic 5n modem with a sampling frequency of $1/T = 74$ GHz (i.e., an optical path resolution of 4 mm or depth resolution of 2 mm); the laser operates at a wavelength of 1550 nm. We communicate with the modem over its QSFP-DD electrical interface [Ghiasi 2023] to program the transmit sequence and read out the measured data. The transmitted sequence length, limited by finite modem memory, is set to $\approx 2^{16}$ symbols, providing a maximum unambiguous range of approximately 130 meters. We use an Erbium-doped fiber amplifier (EDFA [Becker et al. 1999]) to boost the power of the emitted laser light from 1 mW up to a maximum of 100 mW; all experiments use an eye-safe transmit laser power of 2 mW unless otherwise specified. Another EDFA is used as a pre-amplifier to boost the power of the received light up to the level expected by the modem ($\approx$1 mW).

5 Experimental Results

We evaluate FWL across different exposure settings, and we compare to alternative modulation schemes and reconstruction techniques. We also demonstrate the approach for recovery of depth and velocity information, imaging through translucent media, imaging under strong ambient light, reconstructing objects with sub-surface scattering, and reconstruction of room-scale scenes.

5.1 Quantitative and Comparative Evaluation

Generalized matched filtering. We compare our joint estimation framework to a straightforward generalization of matched filtering which incorporates the multiple polarization channels of FWL. Specifically, we modify Equation 8 to correlate the transmit and

received symbols sequences across both polarization channels:

$$\Delta^* = \underset{\Delta}{\mathrm{argmax}} \sum_{p,q} \left| \sum_{\Delta=-\infty}^{\infty} \overline{X}_{n-\Delta}[p] Y_n[q] \right|^2, \quad (15)$$

where $p, q \in \{1, 2\}$ index the polarization channels. While this procedure is convenient because it incorporates information across polarization channels, it does not recover Doppler frequency shift, nor does it recover the complex-valued Jones matrices corresponding to reflections.

Comparison to other modulation schemes. We compare FWL to other modulation schemes emulated with the coherent optical modem. Specifically, we compare FWL to dual-polarization phase-only modulation, dual-polarization amplitude-only modulation, and single-polarization phase and amplitude modulation (see supplement for implementation details). The results are shown in Figure 9; FWL recovers depth maps with fewer outliers compared to the other modalities and compared to depth estimation using generalized matched filtering. We assess depth precision by imaging a planar target at distances of roughly 0.5, 1.0, and 1.5 meters and reporting the mean depth error to a plane fitted to each measurement. We find that FWL outperforms other modalities that do not use all the available degrees of freedom for modulating light (Table 1).

Robustness to noise. We evaluate our method using different per-pixel exposure settings (from 0.125 microseconds to 1 microsecond) while maintaining the same transmit power. At shorter exposure settings the quality of depth estimation degrades (Figure 10); however, the proposed approach demonstrates better performance at low signal to noise ratios compared to generalized matched filtering.

5.2 Imaging Demonstrations

Recovering depth and velocity. In Figure 1 we show a scene illuminated by sunlight (through a window) containing static objects and a spinning hemisphere. The radial velocity of the hemisphere is estimated to be a maximum of 25 m/s, which agrees with estimates produced using a high speed camera (see supplement). Depth is recovered despite the presence of strong ambient light from the sun, which illuminates the scene through the window (see Figure 11 (e)).

Challenging materials and room-sized scene. In Figure 11 (a), we demonstrate reconstructing objects with subsurface scattering and detailed geometry and provide a comparison to generalized matched filtering using optical powers of 2 mW and 80 mW. In each case the qualitative results in Figure 11 show robust reconstruction of the scene geometry. We see similar trends for a larger room-sized scene.

Imaging through translucent media. In Figure 11 (c), we show an example scene where a statuette is placed behind a translucent surface. We recover both surfaces by estimating the time delays of the two Jones matrices with the greatest Frobenius norms.

Strong ambient light. We reconstruct the geometry of a tungsten halide light bulb while it is turned on (Figure 11 (d)). While the bulb's spectrum includes the 1550 nm wavelength used for FWL, the system is extremely robust to ambient light due to the homodyne detection scheme.

6 Discussion

Coherent optical modems are a promising solution to make coherent lidar more accessible to researchers and practitioners. Still, some barriers remain to widespread adoption of this technology. For example, using fast optical modems (operating at tens to hundreds of GHz) requires domain knowledge to program and read out the transmitted and received waveforms. Currently, our hardware interface to the modem requires ~1 second to transfer data to a computer after each exposure; this limits the acquisition speeds of the current system to more than a second per scan point and thus limits the overall scan resolution. In future work we hope to address these obstacles and to demonstrate real-time capture of depth and velocity. FWL may also offer benefits for other imaging scenarios, such as in the presence of scattering media, where sensitivity to motion, depth, and polarization could help to isolate unscattered light.

Acknowledgments

References

Xiao Ai, Richard Nock, John G Rarity, and Naim Dahnoun. 2011. High-resolution random-modulation CW lidar. *Appl. Opt.* 50, 22 (2011), 4478–4488.

Seung-Hwan Baek and Felix Heide. 2021. Polarimetric spatio-temporal light transport probing. *ACM Trans. Graph.* 40, 6 (2021), 1–18.

Seung-Hwan Baek and Felix Heide. 2022. All-photon polarimetric time-of-flight imaging. In *Proc. CVPR*.

Cyrus S Bamji, Swati Mehta, Barry Thompson, Tamer Elkhatib, Stefan Wurster, Onur Akkaya, Andrew Payne, John Godbaz, Mike Fenton, Vijay Rajasekaran, et al. 2018. 1Mpixel 65nm BSI 320MHz demodulated TOF Image sensor with 3μm global shutter pixels and analog binning. In *Proc. ISSCC*.

Sebastian Banzhaf and Christian Waldschmidt. 2021. Phase-code-based modulation for coherent lidar. *IEEE Trans. Veh. Technol.* 70, 10 (2021), 9886–9897.

M Bashkansky, HR Burris, EE Funk, R Mahon, and CI Moore. 2004. RF phase-coded random-modulation LIDAR. *Opt. Commun.* 231, 1-6 (2004), 93–98.

Bernhard Baumann, Conrad W Merkle, Rainer A Leitgeb, Marco Augustin, Andreas Wartak, Michael Pircher, and Christoph K Hitzenberger. 2019. Signal averaging improves signal-to-noise in OCT images: But which approach works best, and when? *Biomed. Opt. Express* 10, 11 (2019), 5755–5775.

Philippe M Becker, Anders A Olsson, and Jay R Simpson. 1999. *Erbium-doped fiber amplifiers: Fundamentals and technology*. Elsevier.

Behnam Behroozpour, Phillip AM Sandborn, Ming C Wu, and Bernhard E Boser. 2017. Lidar system architectures and circuits. *IEEE Commun. Mag.* 55, 10 (2017), 135–142.

Charles A. Brackett. 1990. Dense wavelength division multiplexing networks: Principles and applications. *IEEE J. Sel. Areas Commun.* 8, 6 (1990), 948–964.

Ruixuan Chen, Haowen Shu, Bitao Shen, Lin Chang, Weiqiang Xie, Wenchao Liao, Zihan Tao, John E Bowers, and Xingjun Wang. 2023. Breaking the temporal and frequency congestion of LiDAR by parallel chaos. *Nat. Photonics* 17, 4 (2023), 306–314.

Ian Coddington, William C Swann, Ljerka Nenadovic, and Nathan R Newbury. 2009. Rapid and precise absolute distance measurements at long range. *Nat. Photonics* 3, 6 (2009), 351–356.

G David Forney. 1992. Trellis shaping. *IEEE Trans. Inf. Theory* 38, 2 (1992), 281–300.

Ali Ghiasi. 2023. *QSFP-DD Hardware Specification for QSFP Double Density 8X Pluggable Transceiver*. Technical Report. QSFP-DD MSA. http://www.qsfp-dd.com/wp-content/uploads/2023/09/QSFP-DD-Hardware-Rev7.0.pdf

Ioannis Gkioulekas, Anat Levin, Frédo Durand, and Todd Zickler. 2015. Micron-scale light transport decomposition using interferometry. *ACM Trans. Graph.* 34, 4 (2015), 1–14.

Joseph W Goodman. 2007. *Speckle phenomena in optics: Theory and applications*. Roberts and Company Publishers.

Joseph W Goodman. 2015. *Statistical optics*. John Wiley & Sons.

JP Gordon and H Kogelnik. 2000. PMD fundamentals: Polarization mode dispersion in optical fibers. *Proc. Natl. Acad. Sci. U.S.A.* 97, 9 (2000), 4541–4550.

Felix Heide, Wolfgang Heidrich, Matthias Hullin, and Gordon Wetzstein. 2015. Doppler time-of-flight imaging. *ACM Trans. Graph.* 34, 4, Article 36 (2015), 11 pages.

Sammy W Henderson, Philip Gatt, David Rees, and R Milton Huffaker. 2005. Wind lidar. In *Laser Remote Sensing*. CRC Press, 487–740.

Yunpu Hu, Leo Miyashita, and Masatoshi Ishikawa. 2022. Differential Frequency Heterodyne Time-of-Flight Imaging for Instantaneous Depth and Velocity Estimation. *ACM Trans. Graph.* 42, 1 (2022), 1–13.

Il-Pyeong Hwang and Chang-Hee Lee. 2020. Mutual interferences of a true-random LiDAR with other LiDAR signals. *IEEE Access* 8 (2020), 124123–124133.

Il-Pyeong Hwang, Seok-Jun Yun, and Chang-Hee Lee. 2020. Mutual interferences in frequency-modulated continuous-wave (FMCW) LiDARs. *Optik* 220 (2020), 165109.

Ezra Ip and Joseph M Kahn. 2007. Digital equalization of chromatic dispersion and polarization mode dispersion. *J. Light. Technol.* 25, 8 (2007), 2033–2043.

Ezra Ip, Alan Pak Tao Lau, Daniel JF Barros, and Joseph M Kahn. 2008. Coherent detection in optical fiber systems. *Opt. Express* 16, 2 (2008), 753–791.

Achuta Kadambi, Refael Whyte, Ayush Bhandari, Lee Streeter, Christopher Barsi, Adrian Dorrington, and Ramesh Raskar. 2013. Coded time of flight cameras: sparse deconvolution to address multipath interference and recover time profiles. *ACM Trans. Graph.* 32, 6 (2013), 1–10.

Diederik Kingma and Jimmy Ba. 2015. Adam: A Method for Stochastic Optimization. In *Proc. ICLR*.

Ahmed Kirmani, Dheera Venkatraman, Dongeek Shin, Andrea Colaço, Franco NC Wong, Jeffrey H Shapiro, and Vivek K Goyal. 2014. First-photon imaging. *Science* 343, 6166 (2014), 58–61.

Alankar Kotwal, Anat Levin, and Ioannis Gkioulekas. 2023. Passive Micron-scale Time-of-Flight with Sunlight Interferometry. In *Proc. CVPR*.

PA W Lewis and Gerald S Shedler. 1979. Simulation of nonhomogeneous Poisson processes by thinning. *Nav. Res. Logist. Q.* 26, 3 (1979), 403–413.

Fengqiang Li, Florian Willomitzer, Muralidhar Madabhushi Balaji, Prasanna Rangarajan, and Oliver Cossairt. 2021. Exploiting wavelength diversity for high resolution time-of-flight 3D imaging. *IEEE Trans. Pattern Anal. Mach. Intell.* 43, 7 (2021), 2193–2205.

Fan-Yi Lin and Jia-Ming Liu. 2004. Chaotic lidar. *IEEE J. Sel. Top. Quantum Electron.* 10, 5 (2004), 991–997.

David B Lindell, Matthew O'Toole, and Gordon Wetzstein. 2018. Single-photon 3D imaging with deep sensor fusion. *ACM Trans. Graph.* 37, 4 (2018), 113.

Paul F McManamon. 2015. *Field guide to Lidar*. SPIE Press.

Matthew O'Toole, Felix Heide, David B Lindell, Kai Zang, Steven Diamond, and Gordon Wetzstein. 2017. Reconstructing transient images from single-photon sensors. In *Proc. CVPR*.

Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. 2019. Pytorch: An imperative style, high-performance deep learning library. In *Proc. NeurIPS*.

Christoph Peters, Jonathan Klein, Matthias B Hullin, and Reinhard Klein. 2015. Solving trigonometric moment problems for fast transient imaging. *ACM Trans. Graph.* 34, 6 (2015), 1–11.

John G Proakis and Masoud Salehi. 2008. *Digital communications*. McGraw-Hill.

Joshua Rapp, Yanting Ma, Robin MA Dawson, and Vivek K Goyal. 2021. High-flux single-photon lidar. *Optica* 8, 1 (2021), 30–39.

Joshua Rapp, Julian Tachella, Yoann Altmann, Stephen McLaughlin, and Vivek K Goyal. 2020. Advances in single-photon lidar for autonomous vehicles: Working principles, challenges, and recent advances. *IEEE Signal Process. Mag.* 37, 4 (2020), 62–71.

Kim Roberts, Sik Heng Foo, Michael Moyer, Michael Hubbard, Andrew Sinclair, Jamie Gaudette, and Charles Laperle. 2014. High capacity transport—100G and beyond. *J. Light. Technol.* 33, 3 (2014), 563–578.

Kim Roberts, Qunbi Zhuge, Inder Monga, Sebastien Gareau, and Charles Laperle. 2017. Beyond 100 Gb/s: capacity, flexibility, and network optimization. *J. Opt. Commun. Netw.* 9, 4 (2017), C12–C24.

Callum S Sambridge, James T Spollard, Andrew J Sutton, Kirk McKenzie, and Lyle E Roberts. 2021. Detection statistics for coherent RMCW LiDAR. *Opt. Express* 29, 16 (2021), 25945–25959.

Nicolas Schuhler, Yves Salvadé, Samuel Lévêque, René Dändliker, and Ronald Holzwarth. 2006. Frequency-comb-referenced two-wavelength source for absolute distance measurement. *Opt. Lett.* 31, 21 (2006), 3101–3103.

Sunil Singh and Nar Singh. 2007. Nonlinear effects in optical fibers: origin, management and applications. *Prog. Electromagn. Res.* 73 (2007), 249–275.

David H Sliney and J Mellerio. 2013. *Safety with lasers and other optical sources: A comprehensive handbook*. Springer Science & Business Media.

Nobuo Takeuchi, Hiroshi Baba, Katsumi Sakurai, and Toshiyuki Ueno. 1986. Diode-laser random-modulation cw lidar. *Appl. Opt.* 25, 1 (1986), 63–67.

N Takeuchi, N Sugimoto, Hiroshi Baba, and K Sakurai. 1983. Random modulation CW lidar. *Appl. Opt.* 22, 9 (1983), 1382–1386.

George Turin. 1960. An introduction to matched filters. *IRE Trans. Inf. Theory* 6, 3 (1960), 311–329.

Jesse Zheng. 2005. *Optical frequency-modulated continuous-wave (FMCW) interferometry*. Vol. 107. Springer Science & Business Media.
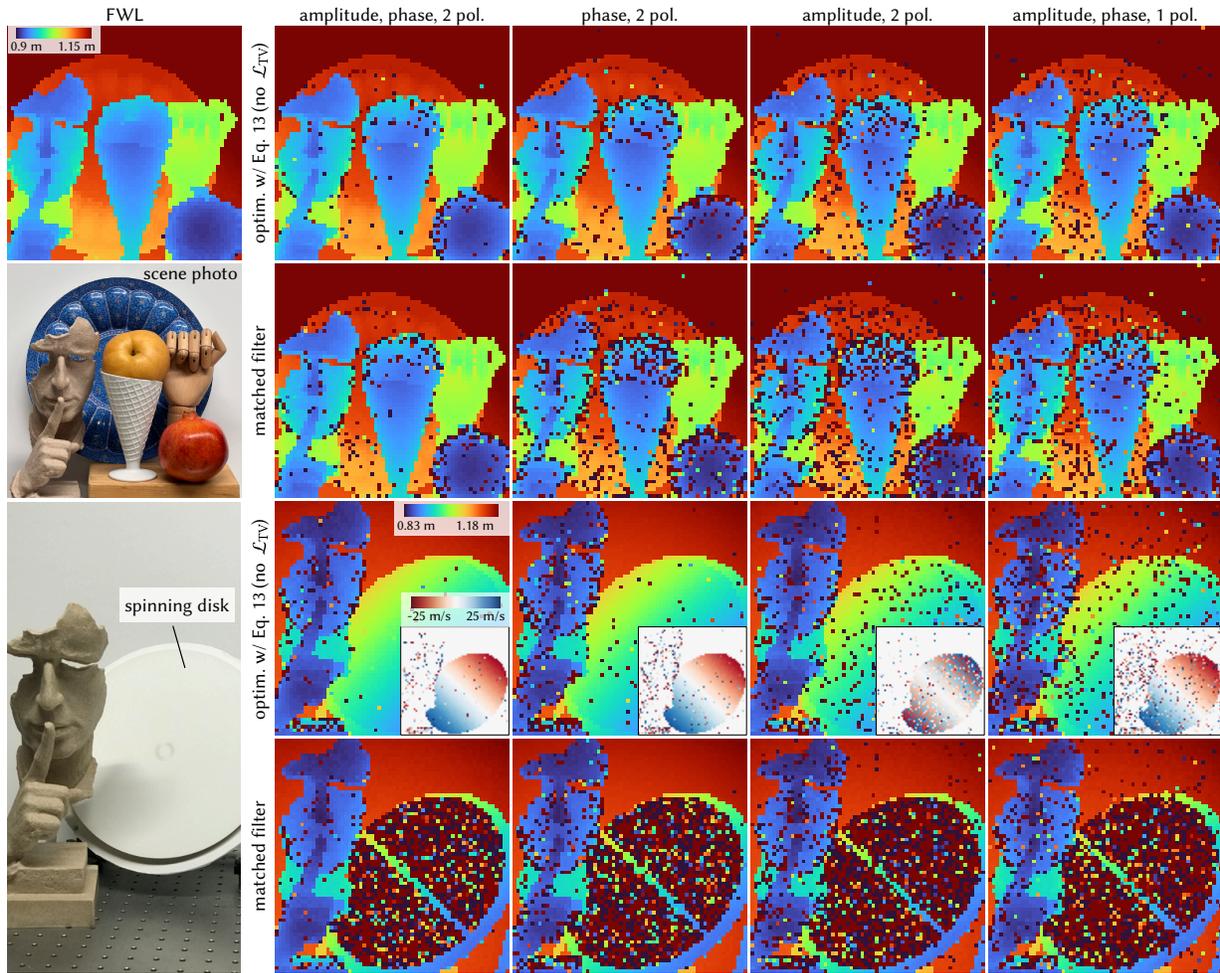
Fig. 9. Comparison between modulation schemes and optimization techniques. **(rows 1, 3)** We compare FWL (amplitude and phase modulation across two polarization channels) to phase-only modulation, amplitude-only modulation, and single-polarization modulation of phase and amplitude without $\mathcal{L}_{TV}$ (see supplement for implementation details). **(rows 2, 4)** Results using generalized matched filtering. FWL optimization based on Equation 13 shows more robust performance compared to other modulation schemes and the generalized matched filter, which also fails to recover accurate depth for Doppler-shifted pixels.
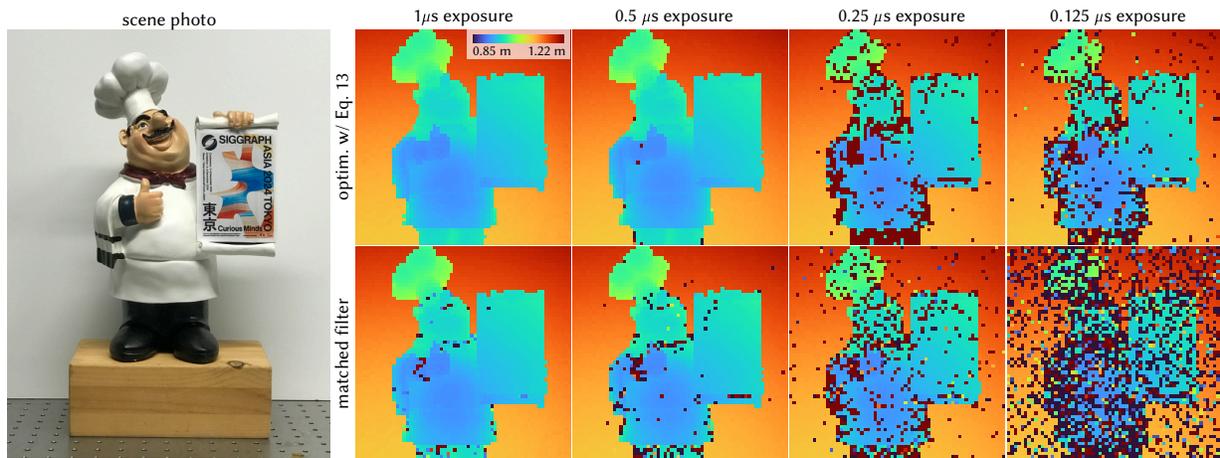

Fig. 10. Optimization using Eq. 13 vs. generalized matched filtering for varying exposure times ranging from $1\mu s$ to $0.125\mu s$. Even with exposures as short as $0.125\mu s$ the proposed technique recovers depth for most pixels while generalized matched filtering fails.

Table 1. Evaluation of depth precision. We scan a planar surface and measure deviation of the measurements to a plane fitted to the surface. We compare to performance using the generalized matched filter; note that we omit the TV regularizer for this evaluation to assess per-pixel precision.

| method	mean depth error (mm) ↓		% of pixels with depth error < 2 mm ↑		% of pixels with depth error < 6 mm ↑	
	joint estimation	gen. matched filter	joint estimation	gen. matched filter	joint estimation	gen. matched filter
FWL	**4.43**	**9.93**	**65.20**	**64.77**	**98.50**	**97.78**
dual-polarization phase	9.31	24.99	57.16	55.40	97.91	94.92
dual-polarization and amplitude	19.08	39.51	58.59	55.72	95.96	91.92
single-polarization phase and amplitude	30.19	46.95	47.85	45.83	93.75	90.46
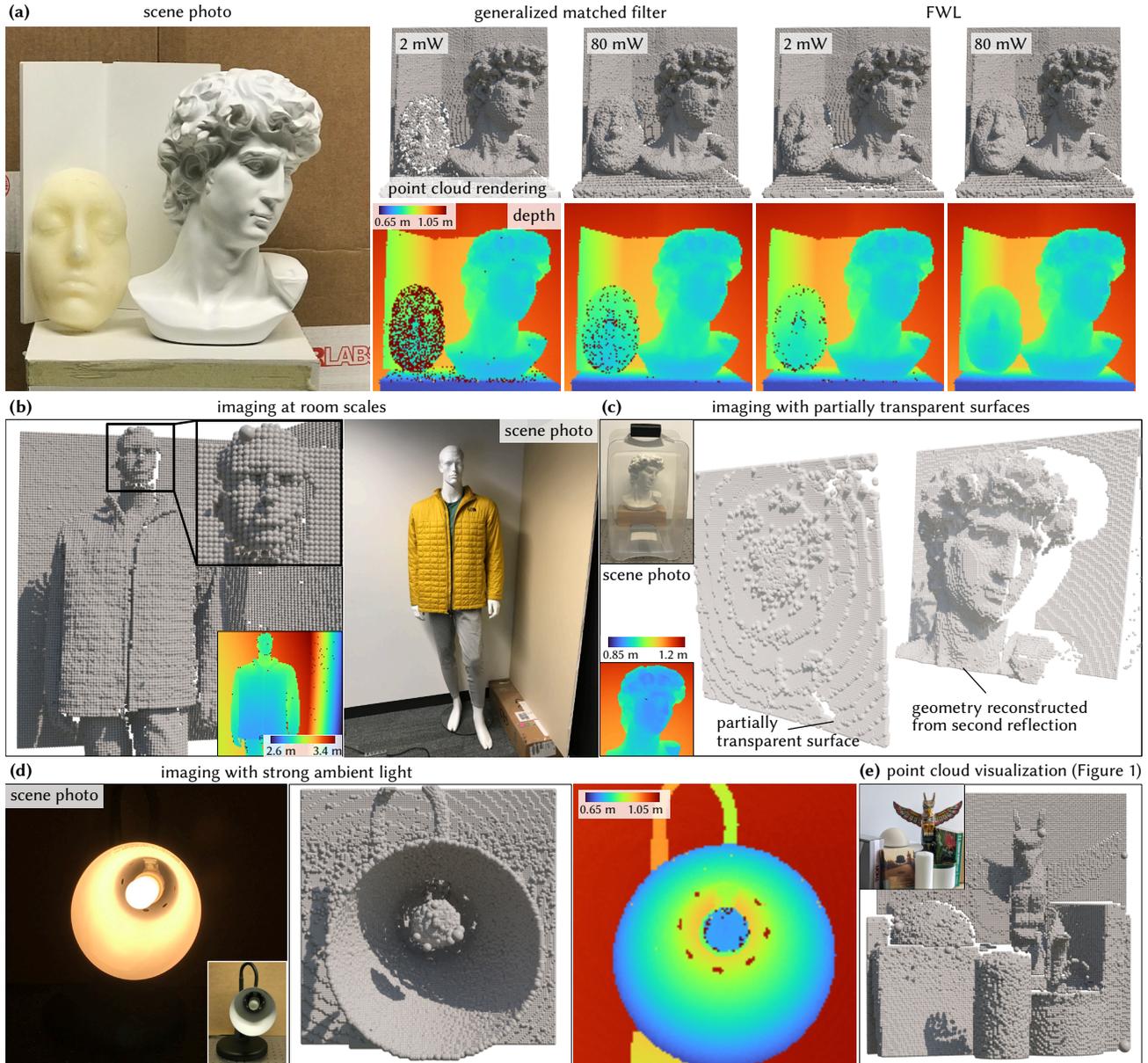
Fig. 11. FWL in challenging imaging scenarios. **(a)** A diffuse statuette and mask made from translucent material are reconstructed with the laser power set to 2 mW and 80 mW. Using the proposed optimization (Equation 13) with the 2 mW measurement recovers a qualitatively similar result to the generalized matched filtering scheme with 80 mW. **(b)** FWL captures a room-sized scene and recovers fine details on the jacket of the mannequin. **(c)** A statuette is placed behind a translucent barrier and the reconstruction captures both the translucent surface as well as the 3D geometry of the statuette. **(d)** FWL scans a tungsten halide light bulb while it is turned on. Although the emission spectrum of the bulb includes the 1550 nm wavelength used by the system, FWL reconstructs both the lamp and the bulb itself. **(e)** A point cloud visualization of the scene in Figure 1. Captures for **(b)**, **(c)**, and **(d)** use a transmit power of 80 mW.