# SDR-ToF: Million Optical Depth Samples per Second Using Software-Defined Radio

RAMCHANDER BHASKARA*, Dartmouth College, USA
DHAWAL SIRIKONDA*, Dartmouth College, USA
OMKAR VENGURLEKAR*, Arizona State University, USA
JUHYEON KIM, Dartmouth College, USA
JOSEPH LAZARRO, Dartmouth College, USA
SUREN JAYASURIYA, Arizona State University, USA
ADITHYA PEDIREDLA, Dartmouth College, USA

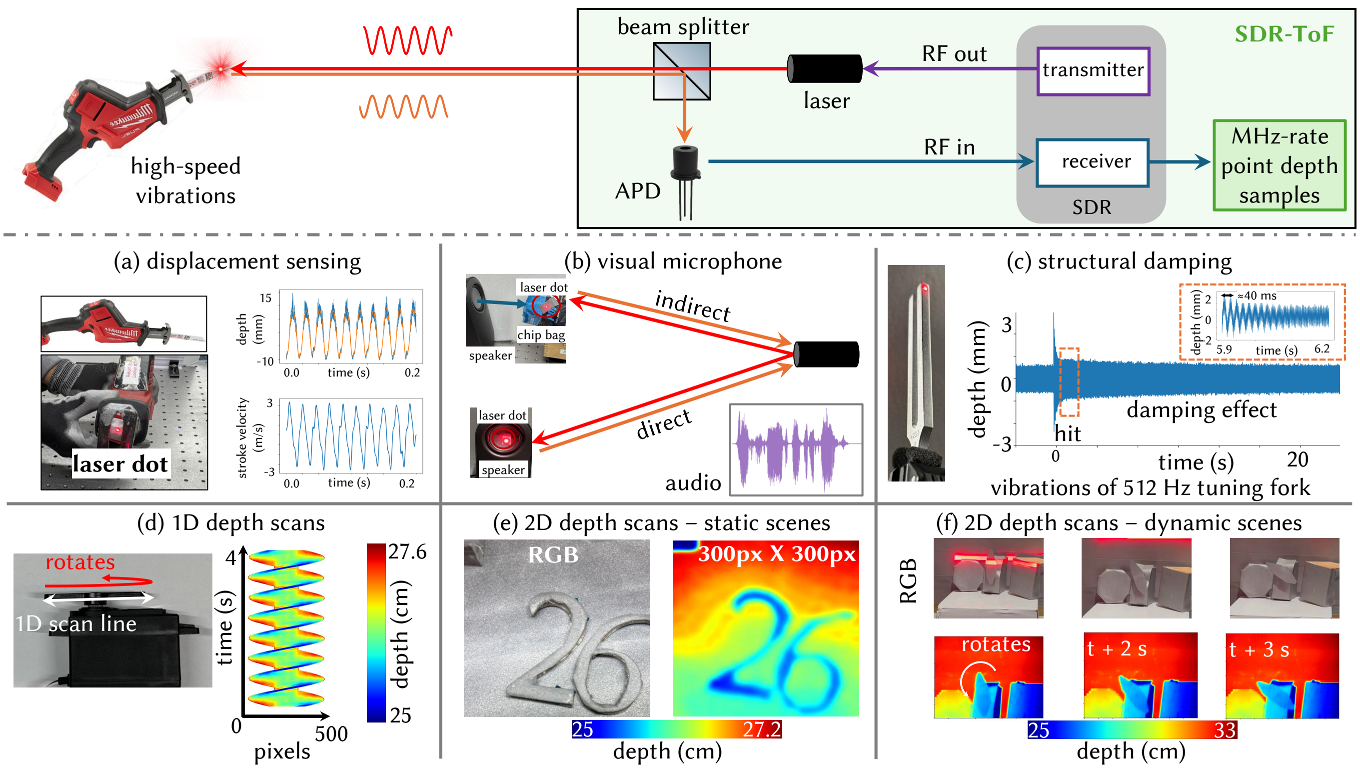


Fig. 1. We introduce SDR-ToF, a time-of-flight (ToF) sensor that uses commodity software-defined radios (SDRs) for RF-coherent optical depth sensing at megahertz rates. SDR-ToF modulates light at near-GHz frequencies to achieve sub-millimeter depth precision while coherently demodulating the received signal to produce a continuous stream of phase measurements, from which metric depth is recovered. We demonstrate applications including: (a) high-precision displacement sensing of mechanical equipment, (b) audio recovery from direct and indirect vibration sources, (c) structural vibration analysis following impact, and (d–f) high-speed 1D and 2D depth scanning using galvanometer-based laser steering. (APD: Avalanche Photo Diode, RF: Radio-Frequency.)

We present a continuous-wave time-of-flight (CW-ToF) sensing system capable of *megahertz-rate depth* measurements using a free-space laser and commodity radio-frequency (RF) communication hardware. Existing high-speed optical displacement and vibration sensing systems typically rely on optical interferometry or specialized heterodyne electronics, making them complex and expensive to deploy. We instead show that *commodity software-defined radios (SDRs)* can serve as a programmable optical sensing backend

*The authors contributed equally to this work.
Project page: https://dartmouth-risc-lab.github.io/sdr-tof/

Authors' Contact Information: Ramchander Bhaskara, Dartmouth College, Hanover, New Hampshire, USA, ramchander.bhaskara@dartmouth.edu; Dhawal Sirikonda, Dartmouth College, Hanover, New Hampshire, USA; Omkar Vengurlekar, Arizona State University, Tempe, Arizona, USA; Juhyeon Kim, Dartmouth College, Hanover, New Hampshire, USA; Joseph Lazarro, Dartmouth College, Hanover, New Hampshire, USA; Suren Jayasuriya, Arizona State University, Tempe, Arizona, USA; Adithya Pediredla, Dartmouth College, Hanover, New Hampshire, USA.




ACM 1557-7368/2026/12-ART400
https://doi.org/10.1111/11111111

by jointly performing laser modulation and RF demodulation of signals measured by a fast photodiode. This architecture enables *direct phase-based depth estimation* at megahertz temporal rates without requiring optical interferometry. Because both the modulation frequency and readout pipeline are fully programmable in software, the system allows flexible trade-offs between depth range, precision, and temporal bandwidth. The resulting sensor supports high-speed depth and vibration measurements, enabling applications such as monitoring rapidly moving mechanical systems, and reconstructing acoustic signals from vibrating surfaces. We experimentally evaluate a prototype system across multiple sensing scenarios and demonstrate *sub-millimeter* depth sensitivity at megahertz sampling rates, while analyzing practical limitations including noise, drift, and hardware constraints. Our results show that commodity RF hardware can enable *low-cost, high-speed optical depth sensing*, opening new opportunities for computational imaging and dynamic scene analysis.

**ACM Reference Format:**
Ramchander Bhaskara, Dhawal Sirikonda, Omkar Vengurlekar, Juhyeon Kim, Joseph Lazarro, Suren Jayasuriya, and Adithya Pediredla. 2026. SDR-ToF: Million Optical Depth Samples per Second Using Software-Defined Radio . *ACM Trans. Graph.* 44, 400, Article 400 (December 2026), 18 pages. https://doi.org/10.1111/11111111

## 1 INTRODUCTION

High-speed measurement of small depth variations is important for capturing dynamic physical phenomena. Applications such as structural vibration analysis, recovery of acoustic waves that insonify surfaces, and tracking of fast-moving objects all require resolving minute displacements at high temporal bandwidths. A key challenge is that vibration amplitudes typically decrease as frequency increases [Rothberg et al. 2017], thus requiring sub-millimeter or micron measurements. Existing sensing systems therefore face a difficult tradeoff between temporal bandwidth, depth precision, and system complexity.

In this work, we present SDR-ToF, a radio-frequency (RF) coherent optical time-of-flight system built using commodity software-defined radio (SDR) hardware. Rather than treating depth sensing as a frame-based imaging problem, we formulate it as a continuous coherent phase measurement problem in the RF domain. Our system intensity-modulates an optical source at RF frequencies and uses a full-duplex SDR transceiver to coherently demodulate the photodetected return signal, producing in-phase and quadrature (I/Q) measurements of optical path delay. This enables megahertz-rate depth sampling with sub-millimeter precision.

The system is conceptually related to continuous-wave time-of-flight (CW-ToF) sensing, where depth is encoded in the phase of modulated intensity of light [Frank et al. 2009; Hansard et al. 2012; Lange 2000]. However, each CW-ToF measurement typically contains a mixture of phase, signal amplitude, ambient illumination, and scene reflectance rather than phase alone. Conventional CW-ToF cameras, therefore, acquire multiple exposures at different phase offsets to isolate phase and estimate depth. Their reliance on frame-based integration and sequential readout limits the effective temporal bandwidth available for rapid, small-amplitude motion [Anthonys 2021; Gupta et al. 2015].

SDR-ToF circumvents this bottleneck by borrowing from coherent RF communication systems. SDR platforms generate and receive coherent RF signals at MHz–GHz rates [Ulversoy 2010], separating phase and amplitude directly via quadrature demodulation. Once the photodiode converts the optical return into an RF signal, the depth sensing problem reduces to coherent demodulation performed via analog electronic circuitry. Coupled with a MHz-rate sampler, this yields a continuous stream of phase measurements at MHz rates, without any frame-based accumulation. In this sense, the SDR acts as an electrical local oscillator and coherent receiver for optical time-of-flight sensing.

Recent works have explored related ideas using optical or electro-optical coherence for point-based depth sensing. Mirdehghan et al. [2024] demonstrates coherent optical sensing by mixing the received optical field with a local reference light beam using optical modem hardware, effectively performing coherent demodulation in the optical domain. Baek et al. [2023] instead use electro-optic modulation (EOM) and RF-coherent optical processing to achieve high phase sensitivity through polarization-coded measurements. These works demonstrate impressive precision using coherent phase sensing architectures. Our goal is complementary: we target high-speed point-based depth sensing with substantially lower optical engineering complexity by moving coherent processing into the RF domain using commodity SDR hardware.

Realizing SDR-ToF requires more than substituting an SDR for RF hardware in conventional CW-ToFs. SDR hardware carries well-known RF front-end imperfections such as DC offset, gain and phase imbalances, and oscillator drift that communications receivers correct using knowledge of transmitted symbols [Schenk 2008]. In our optical ToF sensing, we do not have much tolerance for errors: phase is the measurand, so these RF imperfections introduce systematic depth bias proportional to their magnitude. A one degree phase error at 900 MHz, for instance, produces a ~0.46 mm error in depth.

We therefore design a calibration pipeline around depth estimation rather than symbol recovery. Calibration includes removing the signal-dependent phase distortion that would otherwise set a hard floor on depth accuracy. In particular, we design a digital phase-locked loop to suppress low-frequency oscillator drift that would otherwise manifest as slow depth variation.

Our key contributions include:

**C1** The design and implementation of SDR-ToF, an SDR-based optical time-of-flight prototype that achieves megahertz-rate depth sampling using commodity RF hardware.
**C2** Calibration and digital signal processing techniques for coherent recovery of ToF phase and amplitude.
**C3** Decimation-filtering framework that converts excess sampling bandwidth into measurement precision, along with a quantitative evaluation of its effects on SNR and depth accuracy.
**C4** Experimental validation through optical vibration sensing, velocimetry, and galvo-mirror-based spatial depth scanning (see Fig. 1).

We further analyze the advantages and limitations of the proposed architecture relative to existing approaches and include source code in the supplemental submission.

Table 1. **Comparison with prior work.** We compare prior vibrometry and ToF systems by acquisition characteristics (capture time, cap. time, including exposure and readout; throughput, throu.), system attributes (cost and engineering complexity, eng.), sensing cues (coherence modality, coh.; polarization cues, polar.; modulation frequency, mod. freq.), and sensed quantities (vibration, vibr.; depth; and velocity, vel.). Visual vibrometry methods provide high-throughput temporal sensing but do not recover metric depth, whereas prior ToF approaches recover depth at substantially lower acquisition throughput due to sensing and readout bottlenecks. Our system provides continuous 10 MHz streaming throughput with sub-millimeter depth precision. Please see the supplementary material for a cost comparison.

| **method** | **cap. time** | **throu.** | **cost** | **eng.** | **coh.** | **polar.** | **mod. freq.** | **vibr.** | **depth** | **vel.** |
|---|---|---|---|---|---|---|---|---|---|---|
| Davis et al. [2014], Zhou et al. [2025] | video-rate | video-rate | ≈ $3k | low | NA | ✗ | NA | ✓ | ✗ | ✗ |
| Sheinin et al. [2022] | line-rate | line-rate | ≈ $3k | med. | NA | ✗ | NA | ✓ | ✗ | ✗ |
| Baek et al. [2023] | $\approx 13 \times 10^3$ µs | ≈75 Hz | ≈ $60k | high | RF | ✓ | GHz | ✗ | ≈33 µm @ 13 ms | ✓ |
| Mirdehghan et al. [2024] | 1 µs + 1 s | ≈1 Hz | ≈ $40k | high | optical | ✓ | 74 GHz | ✗ | ≈2 mm @ 1 µs | ✓ |
| **ours** | 0.1 µs | 10 MHz | ≈ $10k | med. | RF | ✗ | ≤2.2 GHz | ✓ | 20 µm @ 10 ms<br>56 µm @ 0.05 ms<br>123 µm @ 1 µs | ✓ |

## 2 RELATED WORK

### 2.1 Time-of-flight depth imaging

Time-of-flight (ToF) imaging is a widely used approach for active depth sensing. Direct ToF (dToF) systems estimate distance by measuring the round-trip travel time of short optical pulses using detectors such as single-photon avalanche photodiodes (SPADs), often using time-correlated single-photon counting (TCSPC) and histogramming to accumulate many pulses for accurate timing [Bronzi et al. 2015; Charbon et al. 2013; O'Connor 2012; Pediredla et al. 2018], increasing system complexity. Increasing depth acquisition speeds for SPADs remains challenging as TCSPC is limited by photon pile-up. In conventional TCSPC, count rates are typically kept to about 1–5% of the laser repetition rate to avoid pile-up distortion [Becker 2005]. Thus, even for a high repetition laser (say 80 MHz) source, the usable detected count rate is only about 0.8 Mcounts/s to 4 Mcounts/s per channel, and precise depth estimates still require accumulating many photons across pulses.

APD LiDARs rely on analog thresholding to estimate pulse round-trip time. Commercial systems such as the Velodyne VLP-16 achieve an aggregate rate of up to 600,000 points per second across 16 channels, corresponding to an effective rate of approximately 37,500 points per second per channel [Ouster, Inc. 2019]. Intel RealSense™ L515 reaches higher aggregate rates by trading dwell time for coverage: L515 sweeps a single pulsed beam to aggregate 23.6 million depth samples per second over a 1024 × 768 frame at 30 FPS [Intel Corporation 2020], so each scene point is sampled 30 times per second at millimeter precision. In contrast, SDR-ToF is a point-based system that generates a continuous stream of 10 million depth points per second and is software programmable. The distinction is not aggregate throughput, which scanned and array systems obtain through spatial multiplexing, but temporal bandwidth at a point.

Indirect ToF (iToF) systems estimate depth from the phase delay of intensity-modulated continuous-wave illumination using correlation sensors such as lock-in pixel arrays and photonic mixer devices (PMDs) [Gupta et al. 2015; Gutierrez-Barragan et al. 2021; Kadambi et al. 2013; Li et al. 2017]. A substantial body of computational imaging work has shown that these correlation sensors serve as programmable measurement backends for transient imaging [Heide et al. 2013, 2014; Peters et al. 2015], radial velocity estimation [Heide et al. 2015], multi-device coordination [Shrestha et al. 2016], and multi-path interference mitigation [Bhandari et al. 2014; Gupta et al. 2015; Kadambi et al. 2016; Naik et al. 2015]. Conceptually, our technique is closest to these techniques. Two key differences distinguish our approach. First, current CW-ToF sensors reach megasample-per-second rates by spatial multiplexing, so each pixel is still sampled at the frame rate, 30 Hz to 60 Hz [Bamji et al. 2014]. SDR-ToF makes the complementary trade, concentrating 10 MHz on a single point. Second, lock-in pixel arrays are typically limited to modulation frequencies of tens of MHz due to on-chip integration constraints [Bamji et al. 2014]. SDR-ToF supports programmable modulation from 50 MHz to 2200 MHz, where higher frequencies yield finer depth precision and lower frequencies extend the unambiguous range.

Frequency-modulated continuous-wave (FMCW) LiDAR systems operate at optical frequencies and estimate range and velocity from beat frequencies generated by chirped optical signals [Hariyama et al. 2018; Iiyama et al. 2011; Passy et al. 1994; Ula et al. 2019; Zhang et al. 2019]. Range estimation requires accumulating a full chirp period to resolve the beat frequency, fundamentally limiting the depth update rate per point. While these systems can achieve sub-millimeter-level precision, they require narrow-linewidth tunable lasers, precise chirp linearization, and coherent detection hardware, resulting in complex optical setups and high system cost. Conversely, SDR-ToF produces a continuous stream of high-precision depth estimates at MHz rate without the accumulation constraints of chirp-based ranging.

### 2.2 High-speed optical sensing

Several works have shown that videos can capture high-frequency scene dynamics without active illumination. Davis et al. [2014] introduced passive visual microphone, demonstrating that subtle vibrations in objects can be used to recover sound. Subsequent work extended these ideas to recovering structural vibrations and modal frequencies from video [Chen et al. 2017; Davis et al. 2015], revealing imperceptible displacements using motion magnification [Wadhwa et al. 2013, 2014], and remote vibration sensing [Lin et al. 2025; Sheinin et al. 2022; Zhang et al. 2023] using additional active illumination techniques. However, these approaches recover motion indirectly from intensity variations, yielding relative rather than metric displacement estimates. In contrast, our system directly measures metric depth at a single spatial point at high temporal resolution, enabling explicit reconstruction of absolute surface displacements.

### 2.3 Laser Doppler vibrometry

Laser Doppler vibrometers (LDVs) measure velocity from the Doppler shift of laser light that is backscattered from a moving surface [Drain 1980; Rembe et al. 2025]. Sub-nanometer displacement sensitivity is achievable, but requires a meticulous interferometric architecture: a reference arm, acousto-optic frequency shifter, and coherent optical detector [Li et al. 2018a,b]. Scanning LDVs extend this to spatial maps through sequential point acquisition [Stanbridge and Ewins 1999]. Wu et al. [2026] uses coherent speckle interferometry, rather than Doppler sensing, to recover dense microscale displacement from single-shot interferograms using two mutually coherent illumination beams. At its core, however, interferometers utilize coherent demodulation of an optical carrier whose phase encodes path-length variation. SDR-ToF performs the demodulation electronically in the RF domain, without an interferometric optical front-end.

### 2.4 SDRs in sensing

SDRs provide a flexible platform for implementing programmable sensing systems by enabling radio-frequency signal generation, demodulation, digitization, and processing entirely in software. SDR-based architectures have been widely used in radar and wireless sensing, where they enable rapid prototyping of modulation, demodulation, and signal processing pipelines without specialized hardware [Guan et al. 2020; Heunis et al. 2011; Hussein et al. 2023; Martone et al. 2025; Rossler et al. 2011]. In optical systems, SDR-like communication hardware has primarily been used in the context of visible-light communication (VLC) systems [Che 2025; Martinek et al. 2019; Wang et al. 2014]. Mahnke [2018] employs an SDR solely as a lock-in receiver with a separate RF synthesizer for laser modulation in frequency modulation spectroscopy. In contrast, our work exploits the inherent transmitter-receiver (Tx-Rx) coherence of a full-duplex SDR transceiver to simultaneously modulate and demodulate an optical signal, enabling high-rate phase-based depth sensing without external synchronization hardware.

### 2.5 RF-optical systems

Gupta et al. [2018]; Gutierrez-Barragan et al. [2019] inspire our work. Their prototypes realize the CW-ToF modulation and demodulation chain (waveform generator, RF amplifier, mixer, low-pass filter, and digitizer) as discrete instruments. SDR-ToF subsumes this chain into a single commodity SDR, where the RF functionalities are all set in software rather than by component selection and cabling. Utilizing SDRs also lifts the RF bandwidth to GHz rates while achieving >10 MHz sampling rates.

Recent work has explored repurposing telecommunication hardware for coherent sensing, which motivates our approach. Mirdehghan et al. [2024] demonstrates full-wavefield point-depth sensing using coherent optical communication modems, achieving approximately 2 mm depth resolution with a 1 µs exposure (with 1 s readout time). However, their architecture requires capturing and retrieving raw waveform data that can reach several gigabytes per measurement, leading to dwell times on the order of seconds. As a result, the system is not well-suited for real-time depth scanning and offers limited flexibility in the sensing pipeline.

Baek et al. [2023] demonstrates GHz-modulated (up to 14.32 GHz) indirect ToF imaging using electro-optical modulators (EOMs). In their design, the intensity-modulated optical signal is optically demodulated to baseband before digitization, and the correlated signal is sampled at 24 kHz. Depth is then estimated through correlation-based phase recovery that requires measurements at multiple phase offsets. This process results in an effective per-point acquisition time of approximately 13 ms. Consequently, despite the high optical modulation frequency, the achievable depth update rate is limited to 75 Hz due to sequential phase stepping and temporal integration.

These prior two works also differ in *where* the received signal is demodulated. Mirdehghan et al. [2024] perform quadrature demodulation in the optical domain using a coherent telecom modem. Baek et al. [2023] mix GHz-rate illumination to baseband optically via EOMs, then recover IQ signals digitally via phase stepping. Traditional CW-ToF cameras similarly reconstruct IQ digitally from four phase-shifted correlation measurements. Our system occupies an unexplored middle ground: an SDR front-end performs quadrature demodulation entirely in the analog electronic domain via RF local oscillators, delivering baseband I/Q samples directly to the ADC — without optical components in the demodulation chain, and without time-multiplexed phase stepping.

Building on this key idea, SDR-ToF leverages a full-duplex SDR to provide programmable modulation and demodulation (up to 2.2 GHz) while directly digitizing the demodulated RF baseband signal at a rate of 10 MHz. Phase is computed directly from in-phase and quadrature samples without requiring sequential phase-offset measurements that reduce measurement bandwidth. This streaming architecture enables MHz-rate phase sampling with sub-millimeter accuracy. Tab. 1 summarizes our work in comparison with prior art.

## 3 TECHNICAL BACKGROUND

In this section, we provide the background necessary to understand the contributions of our paper. Readers familiar with CW-ToF sensing may skip Sec. 3.1, while those familiar with SDR operations of modulation and demodulation may skip Sec. 3.2.

### 3.1 Optical time-of-flight depth sensing

A CW-ToF system estimates scene depth by temporally modulating the emitted optical intensity and measuring the corresponding phase delay in the returned signal. Consider a source with mean radiant intensity $I_0$, modulated at frequency $f$ with modulation contrast $\alpha$. The transmitted radiant intensity is given by

$$I_{\text{tx}}(t) = I_0\left[1 + \alpha \cos(2\pi f t)\right]. \tag{1}$$

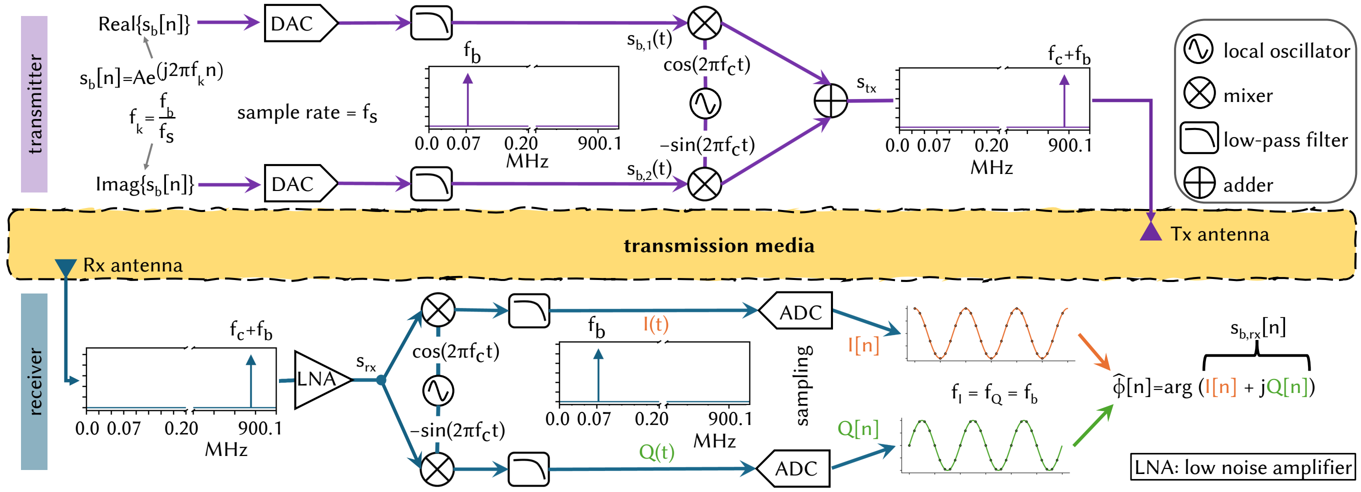


Fig. 2. **Quadrature amplitude modulation and demodulation in an SDR transceiver**. At the transmitter (top), a discrete-time complex baseband signal $s_b[n] = Ae^{j2\pi f_k n}$ is generated, where $f_b$ denotes the baseband frequency, $f_s$ is the sampling rate, and $f_k = f_b/f_s$ is the corresponding normalized digital frequency. The real (Real{$s_b[n]$}) and imaginary (Imag{$s_b[n]$}) components of the complex sinusoid are applied to separate DACs and mixed with a common local oscillator at carrier frequency $f_c \approx 900$ MHz using $\cos(2\pi f_c t)$ and $-\sin(2\pi f_c t)$, respectively. The mixer outputs are summed to form a real RF passband signal centered at $f = f_c + f_b$, which is transmitted through the channel. At the receiver (bottom), the incoming RF signal is coherently downconverted using synchronized in-phase and quadrature local oscillators. After low-pass filtering, the ADC samples the signal at a rate of $f_s$, the discrete-time baseband components I[$n$] and Q[$n$] are recovered, forming the received complex baseband signal $s_{b,\mathrm{rx}}[n]$ = I[$n$] + $j$Q[$n$]. Finally, the received signal's phase is computed as $\hat{\phi}[n] = \arg(s_{b,\mathrm{rx}}[n])$.

After propagating to a scene point at distance $d$ and reflecting back to the sensor, the received signal is an attenuated and time-delayed version of the transmitted waveform,

$$I_{\mathrm{rx}}(t) = \rho I_0 \left[1 + \alpha \cos(2\pi f (t - \tau))\right] + a, \tag{2}$$

where $\rho$ captures geometric and photometric attenuation, $a$ denotes ambient illumination, and

$$\tau = \frac{2d}{c} \tag{3}$$

is the round-trip propagation delay, $c$ is the speed of light.

The propagation delay manifests as a phase shift between the transmitted and received modulation signals,

$$\phi = 2\pi f \tau = \frac{4\pi f d}{c}. \tag{4}$$

Consequently, estimating the phase shift $\phi$ directly yields the optical path length and corresponding scene depth,

$$d = \frac{c\,\phi}{4\pi f}. \tag{5}$$

Since $\phi$ is recoverable only modulo $2\pi$, the depth recovered in Eq. (5) is unambiguous only up to $d_{\max} = c/(2f)$. We discuss this range limitation further in Sec. 6.4.

Conventional CW-ToF systems employ image sensors whose shutter is temporally modulated during the integration interval, enabling correlation between the received optical radiance and a reference modulation waveform. By phase shifting exposure and sampling the resulting correlation measurements, CW-ToFs encode the phase delay between the transmitted and received optical signals.

In a standard CW-ToF demodulation scheme, four phase-shifted correlation measurements are acquired using a reference modulation signal with phase offsets of 0, $\pi/2$, $\pi$, and $3\pi/2$, synchronized with the transmitted illumination. These measurements allow the received signal to be projected onto orthogonal sinusoidal basis functions ($\sin(2\pi f t)$ and $\cos(2\pi f t)$), yielding the in-phase (I) and quadrature (Q) components [Texas Instruments 2014].

The phase of the received signal is then estimated as

$$\phi = \operatorname{atan2}(\mathrm{Q}, \mathrm{I}), \tag{6}$$

from which the scene depth $d$ is obtained [Lange 2000].

The four-phase correlation procedure described above is mathematically equivalent to quadrature demodulation in the communications literature [Lyons 2008]. In CW-ToF, I and Q are formed from discrete phase-shifted correlation measurements, whereas an SDR obtains I/Q by mixing the received RF signal with cosine and sine local oscillator references. Thus, both approaches recover the phase of the modulated return through an equivalent in-phase/quadrature representation, motivating the use of SDRs, which we detail next.

### 3.2 Software-defined radios

Software-defined radios (SDRs) refer to radio systems in which most physical-layer functions are implemented in software rather than hardware [Mitola 2002; Tuttlebee 2002]. A dedicated processing platform performs traditional radio functions such as modulation, demodulation, filtering, and decoding using DSP algorithms. The hardware consists of an RF front end, including antennas, amplifiers, filters, and high-speed analog-to-digital/digital-to-analog converters (ADCs/DACs), which generate and acquire wideband radio-frequency signals. In addition to the RF front end, SDRs contain a field-programmable gate array (FPGA) that performs real-time DSP tasks such as filtering, decimation, and up/downconversion prior to data transfer to the host, offloading time-critical processing from the software running on the host PC.

**Radio transmission.** In communication systems, information signals such as audio or control data occupy a frequency range from DC up to some maximum bandwidth (typically MHz). Such unmodulated signals are called *baseband* signals. Long-range radio transmission, however, requires dedicated RF carriers whose center frequencies typically lie in the GHz range. A baseband signal is transmitted over an RF channel by modulating a high-frequency carrier, shifting its spectral energy from DC to carrier frequency $f_c$, yielding a *passband* signal. Fig. 2 (transmitter) illustrates this upshift in the transmitter section of an SDR transceiver.

**Quadrature amplitude modulation (QAM).** Standard amplitude modulation of a single baseband signal of bandwidth $f_b$ occupies $2f_b$ of RF spectrum, since modulation produces symmetric upper and lower sidebands carrying identical information. Quadrature amplitude modulation (QAM) recovers this factor of two by transmitting two independent baseband signals on the same carrier simultaneously, exploiting the orthogonality of cosine and sine [Stewart et al. 2015].

Specifically, let $s_{b,1}(t)$ and $s_{b,2}(t)$ be two independent real baseband signals, each of bandwidth $f_b$ Hz. Using complex notation, these signals can together be represented as a complex baseband signal $s_b(t)$:

$$s_b(t) = s_{b,1}(t) + j\, s_{b,2}(t), \tag{7}$$

where the real and imaginary parts are referred to as the in-phase (I) and quadrature (Q) components, respectively. Modulating $s_b(t)$ onto a carrier $e^{j2\pi f_c t}$ and transmitting the real part yields

$$s_{\mathrm{tx}}(t) = \mathrm{Re}\left\{s_b(t)\, e^{j2\pi f_c t}\right\} = s_{b,1}(t)\cos(2\pi f_c t) - s_{b,2}(t)\sin(2\pi f_c t), \tag{8}$$

which is the standard QAM passband signal: two independent baseband signals riding on orthogonal carriers within the same $2f_b$ bandwidth. SDRs efficiently implement QAM and Fig. 2 (transmitter) illustrates this process in the transmitter section of the SDR transceiver.

**Quadrature demodulation.** At the receiver (Fig. 2, receiver), the incoming RF signal is coherently downconverted to baseband by multiplying with the complex exponential $e^{-j2\pi f_c t}$, implemented in hardware as simultaneous mixing with in-phase ($\cos 2\pi f_c t$) and quadrature ($\sin 2\pi f_c t$) local-oscillator references.

Assuming ideal propagation i.e., a co-located transmitter and receiver with no path loss and delay, $s_{\mathrm{rx}}(t) = s_{\mathrm{tx}}(t)$ (rx and tx representing receiver and transmitter ends respectively), and substituting Eq. (8),

$$\begin{aligned} s_{\mathrm{rx}}(t)\, e^{-j2\pi f_c t} &= \mathrm{Re}\left\{s_b(t)\, e^{j2\pi f_c t}\right\} e^{-j2\pi f_c t} \\ &= \tfrac{1}{2}\left(s_b(t)\, e^{j2\pi f_c t} + s_b^*(t)\, e^{-j2\pi f_c t}\right) e^{-j2\pi f_c t} \\ &= \tfrac{1}{2}\, s_b(t) + \underbrace{\tfrac{1}{2}\, s_b^*(t)\, e^{-j4\pi f_c t}}_{\text{removed by low-pass filtering}}. \end{aligned} \tag{9}$$

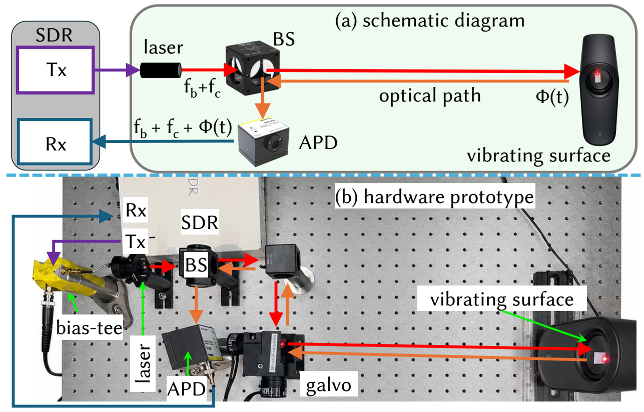


Fig. 3. **Prototype overview.** (a) We use a full-duplex SDR to simultaneously modulate the laser intensity and demodulate the photodetector output on the same board. The vibrating surface introduces a time-varying phase shift, $\Phi(t)$, along the optical path, which the APD captures and the SDR extracts through the demodulation process in Fig. 2 (receiver). (b) Our hardware prototype shows the software-defined radio (SDR), laser, beam-splitter (BS), bias-tee, APD receiver, and vibrating surface used to close the sensing loop, all connected with standard RF cables.

Low-pass filtering (LPF) suppresses the conjugate image term centered at $-2f_c$, recovering

$$s_{b,\mathrm{rx}}(t) = \tfrac{1}{2}\, s_b(t) = \tfrac{1}{2}\left[s_{b,1}(t) + j\, s_{b,2}(t)\right], \tag{10}$$

from which $s_{b,1}(t)$ and $s_{b,2}(t)$ are read off as the real and imaginary parts, respectively. The factor of $\frac{1}{2}$ is a fixed amplitude scaling absorbed into the receiver gain.

In the frequency domain interpretation, multiplication by $e^{j2\pi f_c t}$ at the transmitter shifts the complex baseband spectrum from DC to $f_c$. Because $s_{\mathrm{tx}}(t)$ is real-valued, its spectrum is Hermitian, producing a conjugate image at $-f_c$. At the receiver, multiplication by $e^{-j2\pi f_c t}$ shifts the desired component back to DC and the image to $-2f_c$; the LPF removes the image, leaving $s_{b,\mathrm{rx}}(t)$.

**Sampling and phase retrieval.** From the quadrature demodulation, the recovered baseband signals are sampled at rate $f_s$ using high speed ADCs, producing discrete-time sequences in the I and Q channels ($\mathrm{I}[n]$ and $\mathrm{Q}[n]$), or equivalently $s_{b,\mathrm{rx}}[n] = \mathrm{I}[n] + j\mathrm{Q}[n]$. An instantaneous phase can be obtained as

$$\hat{\phi}[n] = \mathrm{atan2}(\mathrm{Q}[n], \mathrm{I}[n]) = \arg\{s_{b,\mathrm{rx}}[n]\}. \tag{11}$$

## 4 PROPOSED TECHNIQUE & MEASUREMENT MODEL

Having established the principles of time-of-flight sensing and SDRs in Secs. 3.1 and 3.2, we now describe how these concepts are combined in SDR-ToF.

### 4.1 CW-ToF principle and limitation

CW-ToF systems estimate depth from the phase shift of an intensity-modulated optical signal, as reviewed in Sec. 3.1. Conventional CW-ToF cameras recover this phase using correlation measurements acquired at multiple relative phase offsets, typically requiring a four-phase-offset exposure cycle [Gutierrez-Barragan et al. 2021; Li et al.

2014]. Thus, while CW-ToF provides a compact way to measure phase-based depth, its conventional sensor implementation is not well suited for high-bandwidth phase tracking.

### 4.2 SDR-driven optical modulation

SDR-ToF preserves the same phase-sensing principle as CW-ToF, but replaces the multi-phase exposure cycle with quadrature amplitude modulation and coherent quadrature demodulation. The optical source is modulated at an RF frequency $f$, similar to conventional CW-ToF, but we generate this modulation using an SDR. Here, the SDR synthesizes a complex baseband tone at frequency $f_b$ and upconverts it to an RF carrier $f_c$:

$$s_b(t) = a_{\text{tx}} e^{j2\pi f_b t}, \tag{12}$$

$$s_{\text{tx}}(t) = \text{Re}\{s_b(t)e^{j2\pi f_c t}\} = a_{\text{tx}} \cos(2\pi(f_c + f_b)t) \,. \tag{13}$$

The resulting optical intensity modulation frequency is therefore

$$f = f_c + f_b. \tag{14}$$

A nonzero baseband tone $f_b$ is intentionally chosen so that, after coherent downconversion at $f_c$, the phase-bearing return appears at a known intermediate frequency (IF) rather than directly at DC. Direct-conversion SDR receivers are particularly susceptible to impairments concentrated near DC, including local oscillator (LO) leakage, self-mixing-induced DC offsets, flicker noise, and slow gain or bias drifts [Abidi 2002].

By intentionally offsetting the received signal to a nonzero IF at $f_b$, the desired phase-bearing component is separated from low-frequency spurious content and DC artifacts within the receiver chain. The known frequency offset is subsequently removed digitally through the digital downconversion step described in Eq. (21). This effectively implements a low-IF heterodyne architecture in which the signal of interest is translated away from DC prior to phase extraction.

The use of an SDR is not strictly required for generating the optical modulation: a conventional RF waveform generator could also drive the laser or optical modulator at $f$. The key role of the SDR in SDR-ToF lies in the coherent receive chain. Because the transmit and receive chains share a common frequency reference, they remain coherent, enabling direct software-based IQ demodulation of the returned optical modulation and recovery of its absolute phase. In practice, residual slow phase drift remains due to thermal instabilities in the SDR hardware, a known limitation of commodity SDR platforms. We digitally compensate for this in the calibration step described in Sec. 5.2.

After propagation to the scene and back, the reflected intensity incident on the photodetector contains the same modulation frequency $f = f_c + f_b$, but with a propagation-induced phase shift. Following the CW-ToF model in Sec. 3.1, the received optical intensity can be written as

$$I_{\text{rx}}(t) = \rho I_0 \left[1 + \alpha \cos(2\pi f t + \phi(t))\right] + a, \tag{15}$$

where $\rho$ captures optical attenuation and receiver gain, $\alpha$ is the modulation contrast, $a$ represents ambient illumination or slow background terms, and $\phi(t)$ is the propagation-induced phase of the intensity modulation. For a static target, $\phi$ is converted to depth using Eq. (5); for dynamic targets, changes in $\phi(t)$ encode displacement changes.

### 4.3 SDR-driven quadrature demodulation

The reflected optical signal is detected by an APD, which converts optical intensity variations into an electrical photocurrent. After AC coupling and analog conditioning, the phase-bearing electrical signal can be written as

$$i_{\text{rx}}(t) = A_{\text{rx}} \cos(2\pi f t + \phi(t)) \,, \tag{16}$$

where $A_{\text{rx}} \propto \rho\alpha I_0$ absorbs detector responsivity, optical path loss, and receiver gain.

As illustrated in Fig. 2 (receiver), the SDR coherently demodulates this photocurrent using the same carrier frequency $f_c$ used for transmission. Because the transmit and receive chains share a clock, the receiver local oscillator remains phase-locked to the transmitted carrier [Stewart et al. 2015]. The received electrical signal is mixed against two quadrature reference tones and low-pass filtered:

$$r_{\text{I}}(t) = \text{LPF}[2 i_{\text{rx}}(t) \cos(2\pi f_c t)] \,, \tag{17}$$

$$r_{\text{Q}}(t) = \text{LPF}[-2 i_{\text{rx}}(t) \sin(2\pi f_c t)] \,. \tag{18}$$

Since $f = f_c + f_b$, mixing at $f_c$ shifts the phase-bearing modulation term from $f$ down to the intermediate frequency $f_b$, while high-frequency sum terms are removed by the low-pass filters. The resulting complex baseband signal is

$$s_{b,\text{rx}}(t) = r_{\text{I}}(t) + j r_{\text{Q}}(t) = A e^{j(2\pi f_b t + \phi(t))}, \tag{19}$$

where $A$ absorbs constant gain factors.

After analog-to-digital conversion at sampling rate $f_s$, the received complex baseband samples are

$$s_{b,\text{rx}}[n] = A e^{j\left(2\pi \frac{f_b}{f_s} n + \phi[n]\right)}, \tag{20}$$

where $\phi[n] = \phi(n/f_s)$. A subsequent digital downconversion by the known offset $f_b$ removes the residual carrier:

$$s_\phi[n] = s_{b,\text{rx}}[n]\, e^{-j2\pi \frac{f_b}{f_s} n} = A e^{j\phi[n]}. \tag{21}$$

Because the $f_b$ term is set by the transmitted waveform rather than by propagation, removing it leaves the scene-dependent phase $\phi[n]$.

The phase is then recovered as

$$\hat{\phi}[n] = \arg\{s_\phi[n]\} \,, \tag{22}$$

and depth or displacement is obtained using the CW-ToF phase-to-depth relation in Eq. (5).

**Coherent demodulation** Directly digitizing the received RF photocurrent at the SDRs modulation frequency $f$ would require GHz rate samplers. Coherent demodulation avoids this bottleneck by shifting the phase measurement to a much lower frequency before sampling. Prior systems achieve this in different domains: Mirdehghan et al. [2024] performs coherent mixing optically, using the optical beam as a local-oscillator reference, while Baek et al. [2023] performs RF-domain coherent processing using electro-optic modulation.

SDR-ToF is complementary: it also performs coherent demodulation in the RF domain, but uses commodity SDR hardware rather than specialized electro-optic or coherent optical instrumentation, thereby reducing the optical engineering burden. After the APD

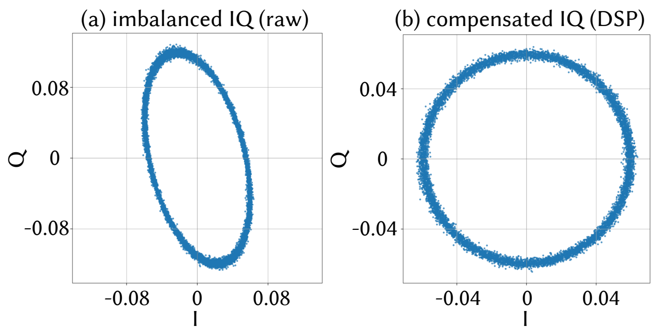


Fig. 4. IQ **imbalance compensation using** 1 ms **capture.** Quadrature demodulation involves mixing received RF signal with local oscillator signals in the orthogonal I and Q channels. Mismatch in phase or amplitude between these channels manifests as an imbalance in the supposedly orthogonal IQ data (sin and cos). (a) amplitude imbalance in IQ resulting in axis-aligned ellipse and phase imbalance causing tilt, along with DC offset causing non zero centered ellipse. (b) blind IQ compensation [Windisch and Fettweis 2004] correcting for offset, gain and phase imbalances in our calibration pipeline. Notice the change in axis limits between (a) and (b); the reduced range in (b) reflects Q-channel gain correction normalizing to the lower I-channel amplitude.

converts the optical return to an RF electrical signal, the remaining pipeline uses standard RF interconnects and software-defined digital downconversion, filtering, and phase extraction. This makes high-bandwidth coherent depth sensing accessible without multi-GS/s digitization or specialized optical demodulation infrastructure, as summarized in Tab. 1.

**Temporal resolution.** The phase is sampled at the SDR ADC rate, $f_s$. In our prototype, $f_s = 10\,\text{MS/s}$, corresponding to one raw phase/depth sample every 0.1 µs—i.e., millions of raw depth samples per second (Contribution **C1**). In applications that do not require MHz measurement bandwidth, we can apply decimation filtering – a programmable bandwidth–precision tradeoff technique that we introduce to improve SNR (more in Sec. 5.2). By choosing the decimation factor and measurement bandwidth, the system can trade temporal bandwidth for depth precision. This enables applications that require higher precision and lower measurement bandwidth. These applications include vibration sensing, structural damping measurements, equipment calibration, and, when combined with beam steering, LiDAR-like scanned depth imaging.

## 5 IMPLEMENTATION DETAILS

We present the benchtop prototype of SDR-ToF in Sec. 5.1, integrating the SDR platform with the optical components. Commodity SDRs natively suffer from RF front-end artifacts and Sec. 5.2 describes the calibration process to mitigate these factors. In Sec. 5.3, we experimentally characterize the depth measurement performance of SDR-ToF.

### 5.1 Hardware Prototype

Fig. 3(a) illustrates the block design of SDR-ToF system while Fig. 3(b) presents an equivalent benchtop prototype. The system indicates signal flow in an RF-to-optical-to-RF path: a full-duplex SDR device generates the RF modulation signal, drives the optical transmitter, and simultaneously receives the photodetector output for quadrature demodulation and phase recovery.

**Transmit chain.** We use a USRP N210 with a WBX daughterboard as the SDR platform.[1] The USRP generates a 900 MHz RF carrier, which is combined with the DC laser bias through a Mini-Circuits ZFBT-4R2G+ bias-tee. This signal directly modulates a free-space Thorlabs LP638P150 laser diode[2] emitting at 638 nm with 10 mW to 15 mW optical power[3]. We use direct laser modulation to support higher RF modulation frequencies than standard laser diode modules.

**Optical path.** The modulated beam passes through a Thorlabs CCM1-BS013 non-polarizing beam splitter, which creates a coaxial transmit–receive path. The beam splitter directs the outgoing beam to the target and redirects the reflected light toward the receiver. This shared optical axis keeps the illuminated point aligned with the receive path and reduces alignment effort. For spatial scanning experiments, we place a Thorlabs GVS012 two-axis galvanometer in the outgoing path to steer the beam across the scene.

**Receive chain.** The returned light is focused by a Thorlabs LA1612-A lens onto a Menlo Systems APD210 avalanche photodetector. The APD converts the reflected intensity-modulated light back into an RF signal, which is attenuated by 30 dB and fed to the USRP receive port. The USRP quadrature downconverts the signal to an intermediate frequency (IF) of $f_b = 10\,\text{kHz}$, digitizes the complex samples, and streams them to the host computer at 10 MS/s for phase estimation. We selected an IF of 10 kHz empirically as a practical operating point away from DC, reducing sensitivity to DC offset, LO leakage, and $1/f$ noise while maintaining stable phase estimation. Higher IF frequencies can increase the phase rotation caused by residual frequency offsets and therefore place greater demands on phase tracking (Secs. A.2 and 5.2).

Overall, the prototype only requires standard RF cabling and coarse optical alignment: the USRP output drives the bias-tee and laser, the beam splitter co-locates the transmit and receive optical paths, and the APD output returns to the USRP. This full-duplex structure lets us recover the target-induced phase variation using the same SDR platform that generates the modulation signal.

The RF bandwidth of our SDR supports up to 2.2 GHz modulation rate. While we can modulate at maximum possible SDR RF carrier frequency (2.2 GHz) to further improve our depth resolution, we chose a modulation frequency of 900 MHz as sensitivity of the APD210 falls off after 900 MHz.

### 5.2 Calibration

Direct-conversion receivers, which demodulate directly to baseband, such as the USRP N210 employed here, exhibit two primary RF front-end impairments: IQ gain and phase imbalance, which introduce unequal scaling across the quadrature channels, and phase noise, which manifests as low-frequency phase drift [Schenk 2008; Tarighat et al. 2005]. Left uncorrected, both corrupt phase-based

[1]The WBX daughterboard supports operation from 50 MHz to 2.2 GHz.
[2]LP638P150 laser diode has a maximum output power of 150 mW.
[3]Class IIIB laser, not eye-safe, so proper lab safety protocols (eyewear, beam enclosure, restricted access) were followed when testing the prototype.

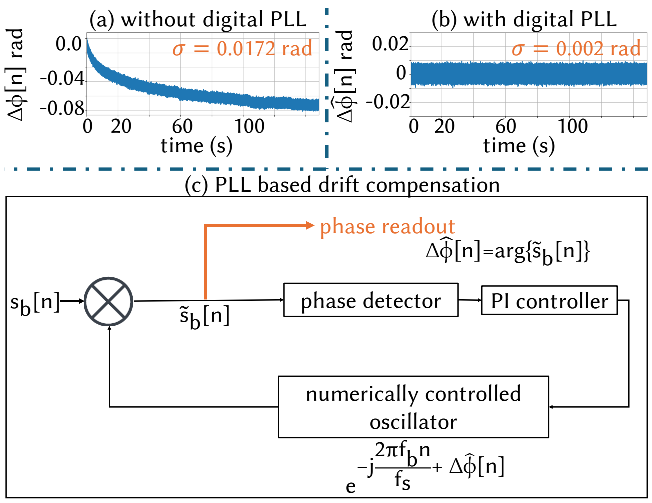


Fig. 5. **Drift compensation using a digital phase-locked loop (PLL).** (a) Raw phase measurements showing low-frequency drift in the SDR receiver. (b) Drift-compensated phase after applying the digital PLL, reducing phase standard deviation ($\sigma$) by approximately 10× over a 140 s window. (c) Digital PLL schematic; the loop tracks slow oscillator drift while preserving higher-frequency phase variations from scene displacement.

depth estimates directly. We address each, along with digital filtering, through the DSP pipeline described below (Contributions **C2** and **C3**).

**IQ imbalance compensation.** The SDR receiver requires the local oscillator to produce sine and cosine waveforms that are precisely orthogonal and of equal amplitude. Synthesizing such waveforms at GHz carrier frequencies in silicon is challenging, and any mismatch manifests as IQ imbalance in the downconverted baseband signal [Abidi 2002]. In the IQ plane, where the in-phase ($\mathrm{I} = \cos(2\pi f_b t)$) and quadrature-phase ($\mathrm{Q} = \sin(2\pi f_b t)$) components of the signal are plotted against each other, a complex sinusoid at a fixed frequency ($f_b$) traces a circle. Imbalance distorts this into an ellipse (Fig. 4a), introducing a systematic phase error that maps directly to depth bias. We apply the blind estimation method of Windisch and Fettweis [2004] to estimate and correct gain and phase imbalance from the statistical properties of the received signal itself, requiring no pilot tones or known reference. Fig. 4(b) shows the corrected result; details are in supplementary.

**Drift compensation.** In SDR-ToF, oscillator phase noise and residual frequency mismatch between the transmit and receive channels introduce a slow time-varying phase trend in the recovered phase signal $\Delta\phi[n]$. Although Tx and Rx share the same onboard reference clock, the analog PLLs operate independently on each channel, introducing residual drift that does not cancel in a differential measurement (phase noise is not common between Tx and Rx).

As shown in Fig. 5(a), this drift produces millimeter-scale baseline errors over long observation windows despite its millihertz rate.

To suppress this, we employ a digital PLL, shown in Fig. 5(c), that continuously tracks and removes the slow phase trend [Gardner 2005; Stewart et al. 2015]. The loop uses a proportional-integral (PI) controller to jointly estimate phase and frequency offset. Referring to Eq. (21), the PLL multiplies $s_b[n]$ with a numerically controlled oscillator (NCO) signal whose phase is continuously corrected by the loop error, yielding the drift-compensated signal $\tilde{s}_b[n]$:

$$\Delta\hat{\phi}[n] = \arg(\tilde{s}_b[n])\,. \tag{23}$$

As shown in Fig. 5(b), the PLL reduces phase standard deviation by approximately 10× over a 140 s observation window. The loop bandwidth of 5 Hz is a tunable parameter: it is set low enough to track only slow drift while leaving higher-frequency phase variations encoding scene displacement intact. Applications requiring tracking of faster dynamics can widen the bandwidth, at the cost of reduced drift rejection; narrowing it improves suppression but increases lock acquisition time. Implementation details are given in the appendix Sec. A.2.

**Decimation filtering.** SDR-ToF acquires IQ samples at 10 MS/s, providing a raw measurement bandwidth of 5 MHz. Most applications require far less: audio recovery demands up to 20 kHz (see Fig. 7), and high-speed actuation sensing typically falls within a few kilohertz (see Fig. 8). The excess bandwidth represents redundant samples that can be exchanged for measurement precision through decimation. As shown in Fig. 6, SNR improves predictably with bandwidth reduction, motivating a deliberate tradeoff between sample rate and depth precision.

We implement a multi-stage decimation filter chain on the complex IQ signal rather than on the extracted phase. There are two practical reasons for this. First, phase is extracted via atan2(Q, I), a nonlinear operation whose output is sensitive to noise in I and Q individually, so filtering IQ first suppresses this noise before the nonlinear step to produce a more accurate phase estimate. Second, deferring phase extraction to the decimated output rate reduces the number of arctangent computations by the factor of decimation, improving the computation speed.

Each decimation stage low-pass filters the signal before downsampling to suppress aliasing. Compared to a moving average decimator, the multi-stage FIR chain provides sharper stopband attenuation and lower per-output-sample arithmetic cost [Harris 2022].

The decimated IQ signal is passed to the digital PLL for drift compensation, which operates at a reduced sampling rate. Following drift correction, a digital downconversion step shifts the baseband frequency $f_b$ to DC, after which phase, and consequently depth, is extracted from the complex baseband samples at a reduced measurement bandwidth.

## 5.3 Characterization

We characterize the noise performance of the proposed sensing system using an optical loopback experiment (Fig. 6(a)), where SDR-ToF measures the displacement of a static target continuously over 150 seconds. From the raw phase samples, we compute the displacement power spectral density (PSD) and report the cumulative root-mean-square (RMS) displacement noise as a function of measurement bandwidth. This representation captures the fundamental precision–bandwidth trade-off of the instrument: increasing measurement bandwidth increases the integrated noise, while operating at lower bandwidth enables improved depth precision through temporal averaging of the high-rate phase samples.

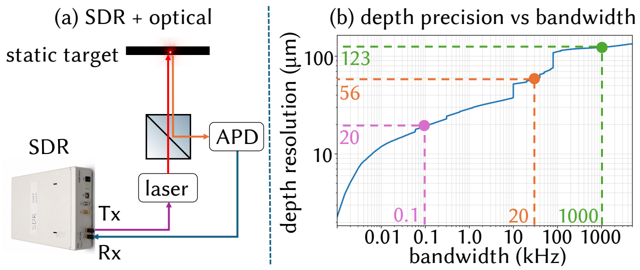


Fig. 6. **Depth resolution of SDR-ToF.** (a) Optical loopback test to measure the intrinsic noise of the electro–optical sensing chain by observing a static target. (b) Depth precision as a function of measurement bandwidth, computed as the cumulative RMS displacement. As bandwidth increases, noise at all frequencies accumulate. The dashed lines indicate an RMS depth noise of ≈20 µm, ≈56 µm, and ≈123 µm at 100 Hz, 20 kHz, and 1 MHz of measurement bandwidths, respectively. Step-like increments in the curve correspond to narrowband spectral spurs in the system.

The resulting curve (Fig. 6(b)) shows that achievable depth precision degrades gradually with increasing measurement bandwidth. Several step-like increments correspond to narrowband spectral spurs in the system, including oscillator leakage near 10 kHz and its harmonics (e.g., 80 kHz), as well as environmental interference near 60 Hz and 300 Hz. Refer to Sec. A.1 for details.

## 6 EXPERIMENTAL EVALUATION

This section presents applications of our SDR-ToF prototype for sensing audio frequencies (Sec. 6.1), optical velocimetry (Sec. 6.2), galvo-based spatial depth scanning (Sec. 6.3), and phase unwrapping for extended-range depth recovery (Sec. 6.4).

The phenomena of interest in these experiments, audio and mechanical vibrations, occupy frequencies well below the 10 MHz depth sampling rate of the system. Audio recovery requires at most 20 kHz of measurement bandwidth, and mechanical motions of interest are typically far below this. Retaining the full sampling bandwidth admits high-frequency noise into the measurements. We therefore apply a multi-stage decimation filter to the IQ samples, reducing the sampling rate from 10 MS/s to 156.25 kS/s through successive decimation stages of factors [8, 4, 2], giving a total decimation factor of 64. Each stage low-pass filters and downsamples the IQ signal, suppressing out-of-band noise and concentrating the available SNR into the retained bandwidth. The achievable depth resolution as a function of measurement bandwidth is presented in Fig. 6. The target rate of 156.25 kS/s provides approximately 8× oversampling of the 20 kHz band, ensuring faithful reconstruction of the highest frequency of interest. A final FIR filter with a 20 kHz cutoff is then applied to suppress residual out-of-band noise before phase is computed at the decimated rate. The complete filter chain is described in Sec. A.1.

### 6.1 Visual microphone: sensing audio frequencies

Surface vibrations appear as fast temporal variations in the recovered depth signal. Tracking these variations reconstructs the acoustic waveform of the vibrating surface, analogous to a microphone. Instead of converting sound pressure to voltage, the system recovers audible frequencies from optical path length changes, which we refer to as a *visual microphone* (coined by Davis et al. [2014]). Unlike speckle-based approaches such as Sheinin et al. [2022], reconstruction arises directly from depth measurements. In communication terms, the surface displacement acts as a baseband signal modulating the RF carrier.

Fig. 7 summarizes three experiments.

**(a) Tuning fork.** A mechanically excited tuning fork provides a controlled monotone vibration. The recovered displacement shows exponential decay from mechanical damping. The spectrum reveals a dominant fundamental at 426 Hz together with higher-order clang mode at 2.585 kHz (second vibrational mode [Russell 2000]), whose excitation varies with the strike location.

**(b) Speaker diaphragm.** A speaker driven by audio signals provides broadband excitation. The recovered displacement follows the diaphragm motion, and the measured spectrum closely matches the input audio, demonstrating non-contact optical audio reconstruction.

**(c) Indirect acoustic reconstruction.** Finally, we measure a secondary surface excited indirectly by sound from a single speaker. Despite weaker vibration amplitude, dominant spectral components of the driving audio remain identifiable, demonstrating sensitivity to indirectly induced motion.

### 6.2 Optical velocimetry

We next evaluate SDR-ToF for large-displacement kinematic tracking under realistic mechanical motion. Because SDR-ToF directly recovers displacement from high-rate phase measurements, the measured depth signal can be numerically differentiated with high-accuracy to estimate instantaneous velocity, enabling optical velocimetry without heterodyne or Doppler-based measurement systems.

We first measure a commercial Multitool[4] operating at maximum power. As shown in Fig. 8(a), a laser spot is placed on the tool housing near the blade mount. The recovered displacement waveform (Fig. 8(b)) shows periodic oscillatory motion with a peak cutting displacement of approximately 3.5 mm. The corresponding velocity trace (Fig. 8(c)) captures the non-sinusoidal delivery of cutting speed at the tool tip. The displacement spectrum (Fig. 8(d)) exhibits a dominant peak near 300 Hz, corresponding to 18 000 oscillations/min (OPM), with additional harmonics indicating non-sinusoidal tool kinematics.

Second, we measure a Hackzall[5] reciprocating saw operated at maximum power. The laser spot is placed near the blade mount, as shown in Fig. 8(e). The recovered displacement signal (Fig. 8(f)) captures the full reciprocating stroke, with an amplitude consistent with the manufacturer-specified stroke length of approximately 22 mm. Differentiating the displacement yields the instantaneous tip velocity (Fig. 8(g)). The depth spectrum (Fig. 8(h)) shows a dominant peak near 50 Hz, consistent with the specified operating speed of 3000 strokes/min, and higher harmonics arising from crank-driven, non-sinusoidal motion.

[4] Milwaukee M18 FUEL™ Oscillating Multi-Tool (accessed May-12-2026)
[5] Milwaukee M18 FUEL™ Hackzall® (accessed May-12-2026)

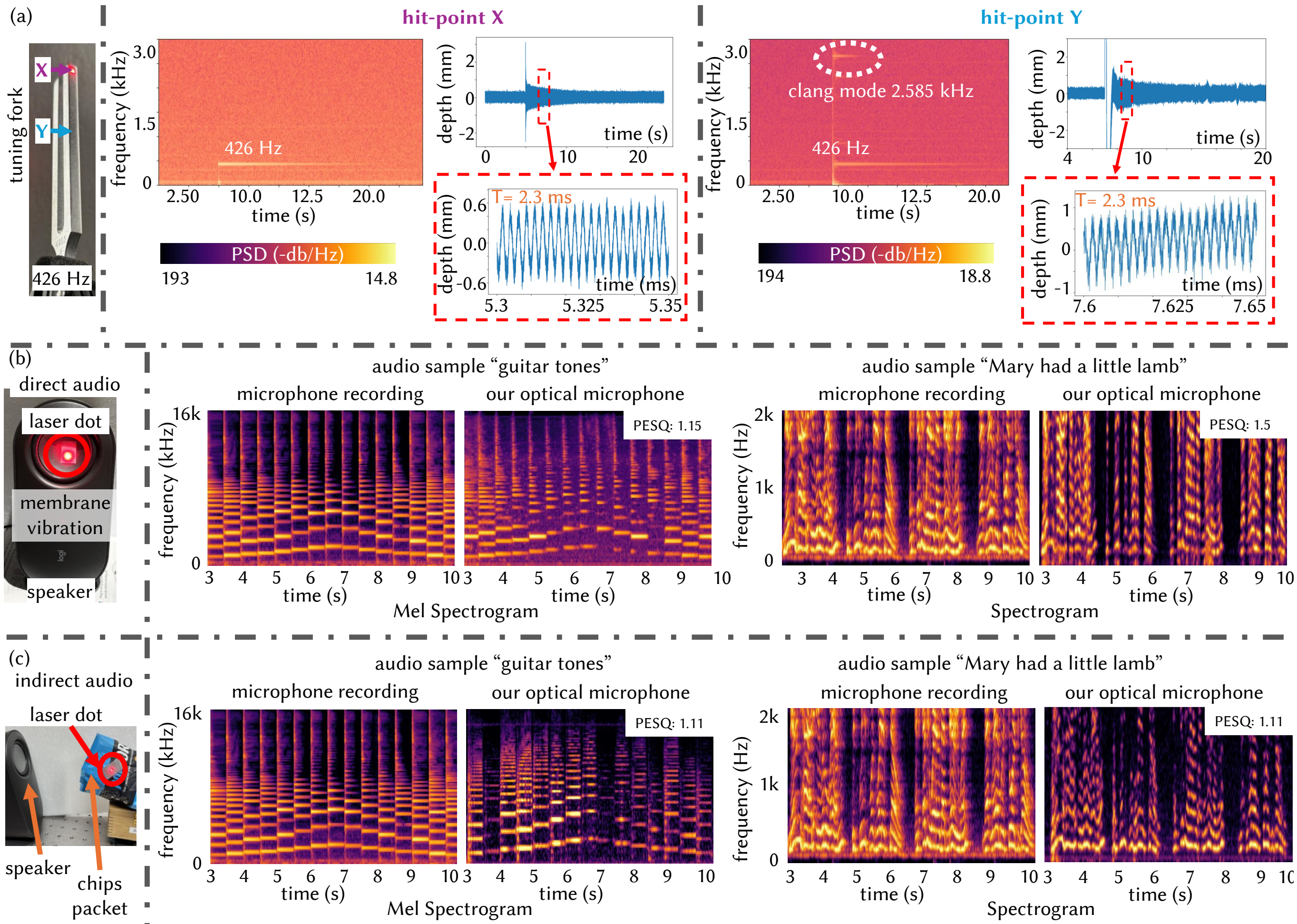


Fig. 7. **Vibration sensing and visual microphone.** (a) Tuning fork excited at the tip (**X**) and middle (**Y**). Both measurements recover the fundamental frequency (426 Hz), while striking the middle additionally excites a clang mode at 2.585 kHz [Russell 2000]. (b) Audio reconstruction from speaker-driven vibrations (guitar mel-spectrogram and spectrogram of spoken verse), similar to Sheinin et al. [2022]. (c) We also demonstrate indirect audio recovery from sound reflected off a chip bag. As in Sheinin et al. [2022], we compare spectral structure relative to the microphone recordings and report Perceptual Evaluation of Speech Quality (PESQ) metric [Rix et al. 2001] for both the capture scenarios. Audio examples are included in the supplementary material.

Note that in Fig. 8(b) and (f), we applied a lowpass Butterworth filter with a cutoff frequency of 2 kHz to filter the depth signal. Fig. 8(c) and (g) indicate the velocities computed from corresponding filtered depth signals.

## 6.3 Scanning with a galvo system

Figs. 9 to 11 show depth imaging with a Thorlabs GVS012 two-axis galvo scanner in three scan modes: 1D, static 2D, and dynamic 2D. Because SDR-ToF measures depth only at a single spot, we scan the beam with sinusoidal drives on the two galvo axes, forming a Lissajous pattern (commonly performed with galvo imaging systems [Liu et al. 2021; Tilmon et al. 2020]) . We reconstruct depth by splatting each measurement $z_i$ at scan position $\mathbf{x}_i$ onto the image grid with a Gaussian weight $w_i(\mathbf{x}) = \exp(-\|\mathbf{x} - \mathbf{x}_i\|^2/2\sigma^2)$, giving $D(\mathbf{x}) = \sum_i w_i(\mathbf{x})z_i / \sum_i w_i(\mathbf{x})$, where $\sigma$ sets the splat size. As the Lissajous trajectories do not fully cover the pixel grid, we use a Gaussian splatting kernel with 2 px standard deviation to spread each sample's contribution and reduce gaps.

**(a) Dynamic 1D spatial scan.** We measure a fan blade rotating at 70 rpm using a one-dimensional scan line aligned with the blade, with the second galvo axis switched off (Fig. 9). The galvo performs a sinusoidal scan at 200 Hz. Plotting the recovered depth along the scan line over time produces a spiral pattern showing blade rotation during scans.

**(b) Static 2D spatial scan.** We use the 2D scan to image a static target, as shown in Fig. 1(e) and Fig. 10. The beam follows a Lissajous trajectory with 100 Hz horizontal and 1 Hz vertical drive frequencies; the vertical sweep gives an effective frame rate of 2 fps. We recover the depth of the bulbasaur-figurine (Fig. 1(e)) on a 300 px×300 px grid. At a sampling rate of 10 MS/s, this gives roughly 56 samples per

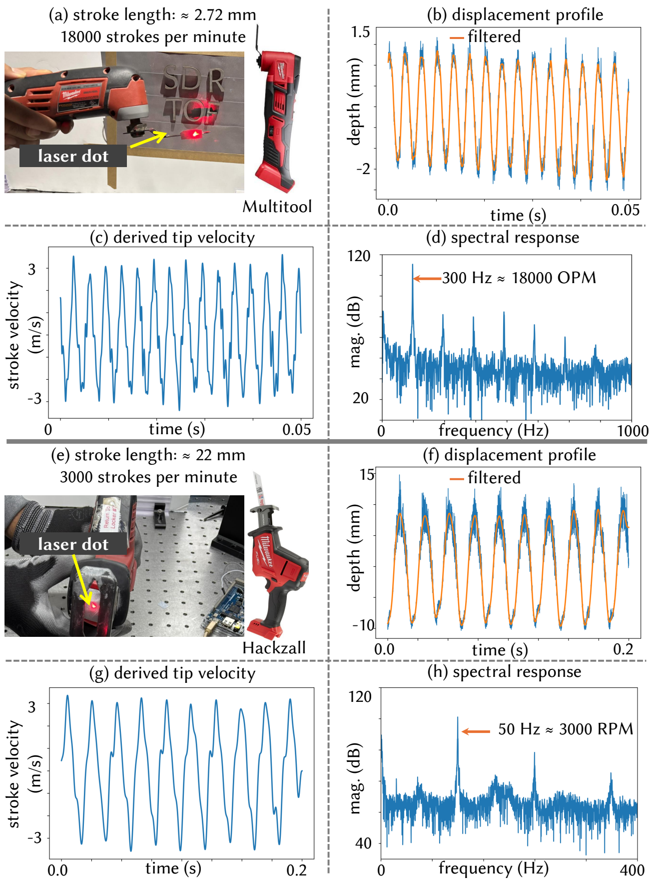


Fig. 8. **High-speed displacement sensing and optical velocimetry of commercial cutting tools.** (a) Laser spot placement on a Milwaukee Multi-Tool near the blade mount with a reflective marker. (b) The recovered depth profile shows periodic oscillatory cutting motion with a peak displacement of approximately 3.5 mm. (c) Instantaneous tip velocity is obtained by differentiating the displacement signal. (d) The displacement spectrum shows a dominant peak near 300 Hz, corresponding to 18 000 oscillations/min (OPM), with higher harmonics from non-sinusoidal tool motion. (e) Laser spot placement on a Milwaukee Hackzall reciprocating saw near the blade mount. (f) The recovered depth profile shows periodic reciprocating motion with a stroke length of approximately 22 mm. (g) Instantaneous tip velocity is obtained by differentiating the displacement signal. (h) The depth spectrum shows a dominant peak near 50 Hz, consistent with an operating speed of 3000 strokes/min, with harmonics from crank-driven motion.

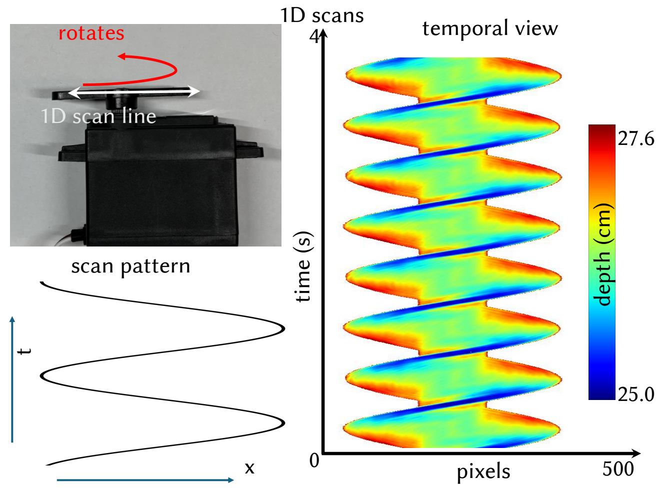


Fig. 9. **Temporal view of one-dimensional scans of a rotating blade**. A scan line intersects a fan blade rotating at 70 rpm. Plotting depth over time produces a spiral structure with a period of ≈0.43 s as expected. We attach a retroreflective marker across the blade length.

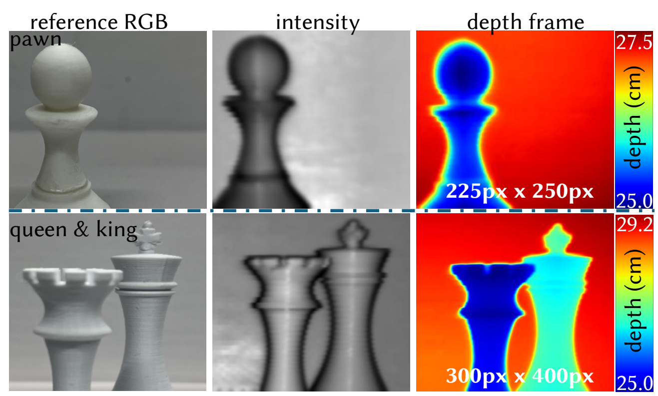


Fig. 10. **Spatial depth scanning of static scenes using a 2D galvo system.** Because SDR-ToF is a point-based sensor, we use a 2D galvo system to scan the scene with a sinusoidal Lissajous trajectory at 100 Hz horizontally and 1 Hz vertically. Top: reference RGB image, recovered intensity, and depth frames for a chess pawn placed in front of a planar background. Bottom: RGB reference, recovered intensity and depth frames for a scene with the queen placed in front of the king. The recovered intensity and depth maps show the respective measurements arranged in a 2D grid. We obtain depth measurements from atan2(Q, I) and intensity measurements from |IQ|.

pixel, or an effective exposure time of 5.6 µs. We also scan two chess-piece scenes against a planar background using the same Lissajous pattern (Fig. 10). For the pawn scene, we recover a 225 px×250 px depth map over a 4.5 cm depth range, with roughly 89 samples per pixel, or 8.9 µs effective exposure. For the queen placed in front of the king, we recover a 300 px×400 px depth map over a 2.2 cm depth range, with roughly 42 samples per pixel, or 4.2 µs effective exposure.

**(c) Dynamic 2D scan.** Fig. 11 shows dynamic reconstruction of a scene with a rotating fan and other static objects, using a Lissajous scan with 100 Hz horizontal and 2 Hz vertical frequencies, producing frames at 4 fps. Depth samples are splatted into image space to form a 250 px×85 px reconstruction. Approximately 118 samples contribute to each pixel, corresponding to an effective exposure time of about 11.8 µs.

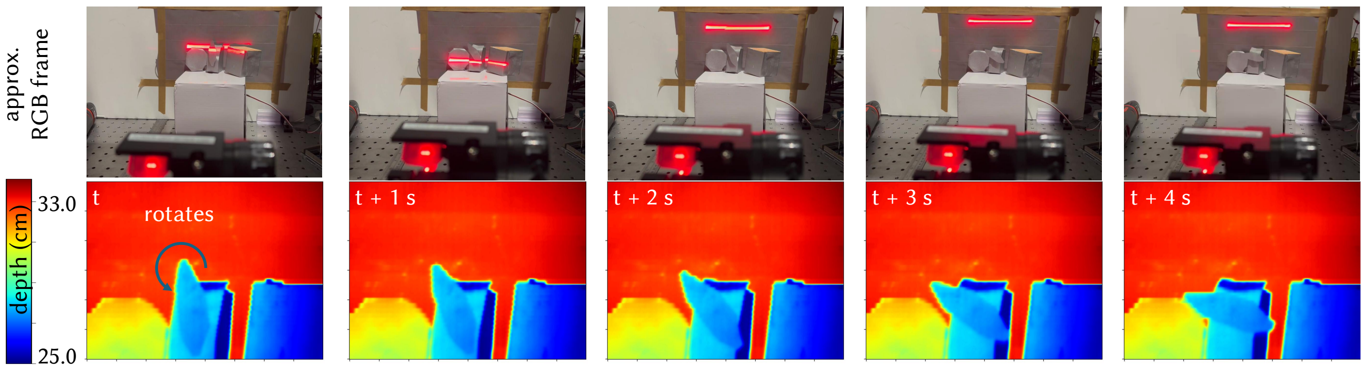


Fig. 11. **Spatial depth scanning of dynamic scenes using a galvo system.** A 250 px×85 px depth reconstruction of a rotating fan with static foreground objects (RGB shown for reference). Dense Lissajous scanning reconstructs the dynamic scene at 4 fps. The top row shows RGB frames from an approximate viewpoint corresponding to the depth scan, while the bottom row shows the reconstructed depth frames in mm. An illustration of the scanning process is provided in the supplementary material.

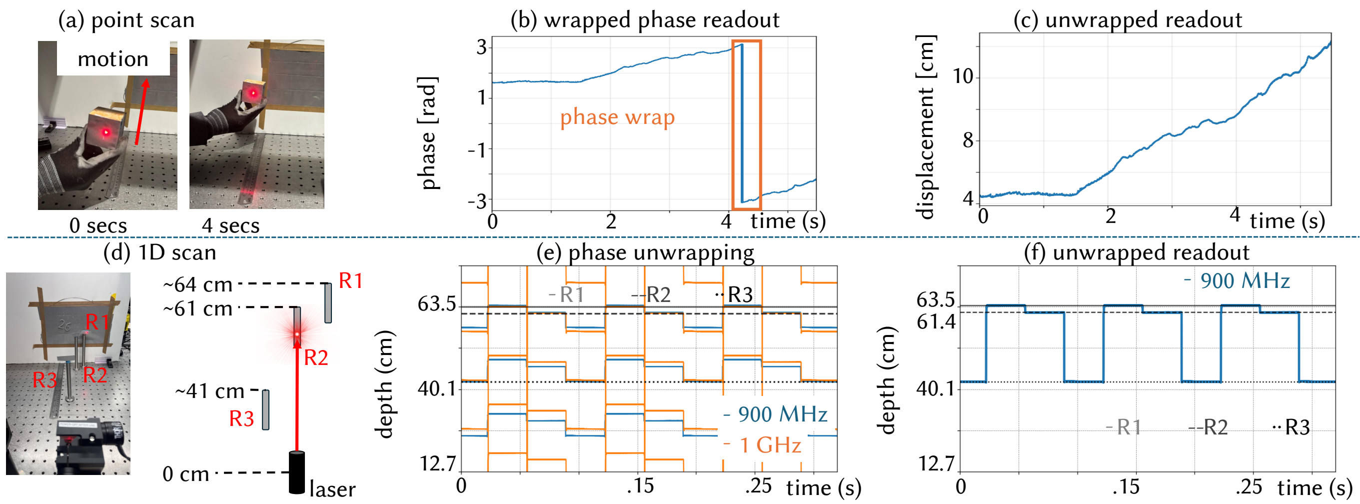


Fig. 12. **Phase unwrapping for depth estimation.** (a) Experimental point-scan demonstrating continuous motion of a target. (b) Wrapped phase measurement obtained from the demodulated signal, where the phase is confined to the principal interval $[-\pi, \pi]$. As the displacement exceeds one phase period, the measured phase abruptly wraps by $2\pi$. (c) Temporally unwrapped phase trajectory obtained by enforcing phase continuity, enabling recovery of the true displacement beyond the wrap boundary. (d) Multi-depth 1D scan consisting of three reflectors (R1–R3) placed at approximately 64 cm, 61 cm, and 41 cm (25 in, 24 in, and 16 in) from the sensor. (e) Candidate distance solutions obtained from two modulation frequencies: 900 MHz and 1000 MHz, with respective unambiguous depth ranges: 16.7 cm and 15 cm. Due to the periodicity of phase measurements, each frequency individually produces multiple ambiguous depth estimates. (f) Final depth reconstruction obtained by identifying the coincident solutions across the two frequencies, thereby resolving the phase ambiguity and recovering the true object depths.

### 6.4 Phase unwrapping

SDR-ToF measures depth from the phase of an intensity-modulated optical signal. Because phase measurements are periodic and constrained to the principal interval $[-\pi, \pi]$, the measured phase corresponds to distance only modulo the ambiguity range of the modulation frequency. For a modulation frequency of 900 MHz, this ambiguity range is 16.65 cm, meaning that depths separated by integer multiples of this distance produce identical phase measurements.

To recover distances beyond this limit, the phase must be unwrapped. Fig. 12(a–c) illustrates a point-scan experiment in which continuous target motion enables temporal phase unwrapping. Although the measured phase exhibits abrupt discontinuities when crossing the $\pm\pi$ boundary (Fig. 12(b)), enforcing phase continuity recovers the true displacement trajectory (Fig. 12(c)).

In static scenes where temporal continuity is unavailable, ambiguity can instead be resolved using multi-frequency measurements, a strategy widely used in CW-ToF sensing [Gupta et al. 2015; Järemo Lawin et al. 2016; Vasisht et al. 2016]. We demonstrate this approach using two modulation frequencies, 900 MHz and 1000 MHz, while scanning three reflectors located at different depths (Fig. 12(d)). Because each frequency produces periodic distance estimates, the measured phase corresponds to multiple possible depth candidates (Fig. 12(e)). By jointly considering the candidates from both frequencies, the correct depth is obtained at the locations where the candidate sets coincide.

Applying this procedure recovers the true reflector depths at approximately 63.5 cm, 61.5 cm, and 40 cm centimeters for reflectors R1, R2, and R3, respectively, as shown in Fig. 12(f).

## 7 DISCUSSION

***Calibration for absolute depth.*** Although the Tx and Rx synthesizers are locked to a common reference clock, we observe a slow residual phase drift over time (See Fig. 5). This behavior originates from independent analog PLLs on the Tx and Rx chains of the SDR as well as thermal variations. The drift manifests as a low-frequency error that gradually perturbs the recovered depth when below the implemented digital PLL's loop bandwidth. To maintain stable depth estimates, we require a reference target and periodic recalibration, which is a limitation of the current prototype.

***Performance under reduced SNR.*** As with conventional CW-ToF systems, our depth precision degrades under reduced-SNR conditions. Sanmartin-Vich et al. [2024] show that phase-estimation error in CW-ToF increases with background illumination and decreases with the amplitude of the received modulated signal. Thus, low-albedo surfaces, which produce weaker returns, and strong ambient illumination both increase phase uncertainty and consequently depth variance. These effects are fundamental SNR limitations of CW-ToF sensing and are therefore also inherited by our SDR-based implementation.

***Design choices and cost reduction.*** Intensity modulation solutions currently available on the market, such as the Thorlabs LDM56/90, restrict modulation frequencies to approximately 600 MHz. To address this limitation, we used a traditional low-cost DC power supply to bias the laser diode above threshold and an inexpensive Mini-Circuits ZFBT-4R2G+ wideband bias–tee to modulate the intensity waveform. The measured optical power reaching the scene averaged approximately 10 mW to 15 mW. After reflection, the APD received approximately 150 μW of optical power from retroreflective targets and only 10 μW of power from non-retroreflective white targets. This low received power constrains our system to use a photodetector with gain. In this implementation, we used a Menlo Systems APD210, which costs approximately $3000 and represents most expensive component in our system. If we drive the laser at higher power, we can significantly reduce the system cost by using a low-cost photo-diode (Thorlabs DET25A at $270[6]) while enabling a broader range of LiDAR applications.

***Noise modeling and priors.*** Our sensing architecture provides a large temporal measurement bandwidth that can be exploited to improve depth precision through improved noise modeling and signal processing. In this work, we partially leverage this bandwidth by aggregating measurements over short temporal windows to improve estimation accuracy. Developing more accurate statistical models of measurement noise and jointly estimating depth over longer temporal windows could further improve SNR and robustness.

In addition, many practical scenarios provide useful prior information such as scene sparsity, geometric constraints, or known frequency range of vibrations that could be incorporated into the depth estimation process. Leveraging such priors through model-based or learning-based approaches could further improve accuracy without requiring changes to the sensing hardware.

***Signal processing.*** Our choice to stabilize phase drift in DSP rather than in hardware reflects a tradeoff between system cost and component complexity. Alternative approaches include disciplining the USRP's internal clock with an external GPS-disciplined reference clock, or characterizing phase response directly with a vector network analyzer (VNA). External oscillators improve frequency stability but add hardware cost and system complexity. VNAs offer an alternative route to phase characterization: handheld units are relatively inexpensive but limited to low sampling rates (e.g., 48 kHz), while high-bandwidth benchtop VNAs capable of sub-6 GHz operation cost substantially more than our USRP and require offline processing rather than real-time correction. Our digital PLL achieves micron-level depth resolution (Fig. 6) using only the USRP's internal clock, minimizing the hardware footprint needed to reproduce SDR-ToF. Lowering phase noise using external clocks and offline VNA characterization are complementary directions for future work.

***Throughput and latency.*** SDRs generate millions of samples per second, and transferring these samples off-board can become a communication bottleneck. Our experiments use the USRP N210 at 10 MS/s, although its 14-bit ADC samples at a nominal 100 MS/s and its Gigabit Ethernet interface can theoretically stream up to 25 MS/s full-precision I/Q samples. In practice, higher sustained streaming rates were unstable in our prototype and led to packet drops during continuous acquisition. Future work can move phase extraction onto the FPGA to exploit more of the ADC rate while avoiding the host-streaming bottleneck. Signal processing techniques to reconstruct depth from dropped samples is also an exciting direction.

***Noise sources.*** Commercial SDR platforms are designed primarily for wireless communication and introduce several RF noise sources that affect precision phase sensing. As shown in Fig. 6, spurious spectral components appear at 60 Hz and its harmonics, originating from power-line coupling into the RF front-end. Additional narrowband spurs appear near 10 kHz and its harmonics (e.g., 80 kHz), which we attribute to oscillator leakage in the RF synthesizer chain. Our current calibration pipeline (Sec. 5.2) addresses systematic phase biases but does not model or suppress these narrowband interference sources. Developing interference-aware calibration or employing shielded, low-spurious SDR platforms represents a meaningful direction for improving depth precision in future work.

***Spatial resolution and frame rates in 2D scans.*** Two factors limit the spatial resolution of our 2D scans. The collimated beam diameter of approximately 2 mm sets a floor on lateral resolution, and motor vibrations in the galvo scanner introduce depth errors during scanning that are absent in stationary point measurements. Both are hardware limitations of the current prototype rather than

[6] Thorlabs DET25A - a 2 GHz Si free-space photodetector (accessed May-12-2026)

fundamental constraints of the approach. Narrower beam optics would improve lateral resolution directly; replacing the galvo with a faster, lower-vibration steering mirror (e.g. the Optotune scanning mirror) would improve frame rates (10 fps to 20 fps) while also reducing motion-induced depth errors.

## 8 CONCLUSION

Perceiving high-frequency, low-amplitude motion is challenging yet important for many scientific and engineering applications. We presented SDR-ToF, a computational time-of-flight imaging system that leverages a commercially available software-defined radio (SDR) to enable high-rate, phase-based depth sensing. By performing quadrature demodulation at near-GHz carrier frequencies and processing millions of depth samples per second, the system achieves sub-millimeter precision. We quantitatively characterized the system and demonstrated applications in motion sensing, optical velocimetry, and depth scanning, illustrating its ability to recover high-bandwidth displacement signals.

More broadly, SDR-driven computational ToF sensing opens a pathway toward a new class of cost-effective, high-bandwidth imaging systems. Such systems could enable applications including structural monitoring in aerospace and civil infrastructure, acoustic and vibration analysis, surgical and machine-tool calibration, remote speech recovery, and bandwidth-limited 3D sensing.


## Acknowledgments

We used large language models for copyediting and paraphrasing during manuscript preparation.


## References


Asad A Abidi. 2002. Direct-conversion radio transceivers for digital communications. *IEEE Journal of solid-state circuits* 30, 12 (2002), 1399–1410.

Gehan Anthonys. 2021. *Investigation of jitter on full-field amplitude modulated continuous wave time-of-flight range imaging cameras.* Ph. D. Dissertation. The University of Waikato.

Seung-Hwan Baek, Noah Walsh, Ilya Chugunov, Zheng Shi, and Felix Heide. 2023. Centimeter-wave free-space neural time-of-flight imaging. *ACM Transactions on Graphics* 42, 1 (2023), 1–18.

Cyrus S Bamji, Patrick O'Connor, Tamer Elkhatib, Swati Mehta, Barry Thompson, Lawrence A Prather, Dane Snow, Onur Can Akkaya, Andy Daniel, Andrew D Payne, et al. 2014. A 0.13 $\mu$m CMOS system-on-chip for a 512× 424 time-of-flight image sensor with multi-frequency photo-demodulation up to 130 MHz and 2 GS/s ADC. *IEEE Journal of Solid-State Circuits* 50, 1 (2014), 303–319.

Wolfgang Becker. 2005. *Advanced time-correlated single photon counting techniques.* Springer.

Ayush Bhandari, Achuta Kadambi, Refael Whyte, Christopher Barsi, Micha Feigin, Adrian Dorrington, and Ramesh Raskar. 2014. Resolving multipath interference in time-of-flight imaging via modulation frequency diversity and sparse regularization. *Optics letters* 39, 6 (2014), 1705–1708.

Danilo Bronzi, Federica Villa, Simone Tisa, Alberto Tosi, and Franco Zappa. 2015. SPAD figures of merit for photon-counting, photon-timing, and imaging applications: a review. *IEEE Sensors Journal* 16, 1 (2015), 3–12.

Edoardo Charbon, Matt Fishburn, Richard Walker, Robert K Henderson, and Cristiano Niclass. 2013. SPAD-based sensors. In *TOF range-imaging cameras.* Springer, 11–38.

Ming Che. 2025. Software-Defined Visible Light Communication for Internet of Things: A Low-Complexity Approach. In *Telecom*, Vol. 6. MDPI, 31.

Justin G Chen, Abe Davis, Neal Wadhwa, Frédo Durand, William T Freeman, and Oral Büyüköztürk. 2017. Video camera–based vibration measurement for civil infrastructure applications. *Journal of Infrastructure Systems* 23, 3 (2017), B4016013.

Abe Davis, Katherine L Bouman, Justin G Chen, Michael Rubinstein, Fredo Durand, and William T Freeman. 2015. Visual vibrometry: Estimating material properties from small motion in video. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 5335–5343.

Abe Davis, Michael Rubinstein, Neal Wadhwa, Gautham J Mysore, Frédo Durand, and William T Freeman. 2014. The visual microphone: passive recovery of sound from video. *ACM Transactions on Graphics (TOG)* 33, 4 (2014), 1–10.

Leslie E Drain. 1980. The laser Doppler techniques. *Chichester* (1980).

Mario Frank, Matthias Plaue, Holger Rapp, Ullrich Köthe, Bernd Jähne, and Fred A Hamprecht. 2009. Theoretical and experimental error analysis of continuous-wave time-of-flight range cameras. *Optical Engineering* 48, 1 (2009), 013602–013602.

Floyd M Gardner. 2005. *Phaselock techniques.* John Wiley & Sons.

Junfeng Guan, Arun Paidimarri, Alberto Valdes-Garcia, and Bodhisatwa Sadhu. 2020. 3D imaging using mmWave 5G signals. In *2020 IEEE Radio Frequency Integrated Circuits Symposium (RFIC).* IEEE, 147–150.

Mohit Gupta, Shree K Nayar, Matthias B Hullin, and Jaime Martin. 2015. Phasor imaging: A generalization of correlation-based time-of-flight imaging. *ACM Transactions on Graphics (ToG)* 34, 5 (2015), 1–18.

Mohit Gupta, Andreas Velten, Shree K Nayar, and Eric Breitbach. 2018. What are optimal coding functions for time-of-flight imaging? *ACM Transactions on Graphics (TOG)* 37, 2 (2018), 1–18.

Felipe Gutierrez-Barragan, Huaijin Chen, Mohit Gupta, Andreas Velten, and Jinwei Gu. 2021. itof2dtof: A robust and flexible representation for data-driven time-of-flight imaging. *IEEE Transactions on Computational Imaging* 7 (2021), 1205–1214.

Felipe Gutierrez-Barragan, Syed Azer Reza, Andreas Velten, and Mohit Gupta. 2019. Practical Coding Function Design for Time-Of-Flight Imaging.. In *CVPR*, Vol. 1. 2.

Miles Hansard, Seungkyu Lee, Ouk Choi, and Radu Patrice Horaud. 2012. *Time-of-flight cameras: principles, methods and applications.* Springer Science & Business Media.

Tatsuo Hariyama, Phillip AM Sandborn, Masahiro Watanabe, and Ming C Wu. 2018. High-accuracy range-sensing system based on FMCW using low-cost VCSEL. *Optics express* 26, 7 (2018), 9285–9297.

Fredric J Harris. 2022. *Multirate signal processing for communication systems.* River Publishers.

Felix Heide, Wolfgang Heidrich, Matthias Hullin, and Gordon Wetzstein. 2015. Doppler time-of-flight imaging. *ACM Trans. Graph.* 34, 4, Article 36 (July 2015), 11 pages. doi:10.1145/2766953

Felix Heide, Matthias B Hullin, James Gregson, and Wolfgang Heidrich. 2013. Low-budget transient imaging using photonic mixer devices. *ACM Transactions on Graphics (ToG)* 32, 4 (2013), 1–10.

Felix Heide, Lei Xiao, Andreas Kolb, Matthias B Hullin, and Wolfgang Heidrich. 2014. Imaging in scattering media using correlation image sensors and sparse convolutional coding. *Optics express* 22, 21 (2014), 26338–26350.

Sebastiaan Heunis, Yoann Paichard, and Michael Inggs. 2011. Passive radar using a software-defined radio platform and opensource software tools. In *2011 IEEE RadarCon (RADAR).* IEEE, 879–884.

Kamal Hussein, Ahmed SI Amar, Abdelhalim Zekry, Mohamed Abouelatta, Ahmed Khairy Mahmoud, and Mohamed Mabrouk. 2023. Design and implementation of sdr platform for radar applications. In *2023 International Telecommunications Conference (ITC-Egypt).* IEEE, 372–375.

Koichi Iiyama, Shin-ichiro Matsui, Takao Kobayashi, and Takeo Maruyama. 2011. High-resolution FMCW reflectometry using a single-mode vertical-cavity surface-emitting laser. *IEEE Photonics Technology Letters* 23, 11 (2011), 703–705.

Intel Corporation. 2020. *Intel RealSense LiDAR Camera L515 Datasheet.* Technical Report Revision 002. Intel Corporation. https://www.mouser.com/datasheet/2/612/Intel_RealSense_LiDAR_L515_Datasheet_Rev002-1713847.pdf Accessed: 2026-08-21.

Felix Järemo Lawin, Per-Erik Forssén, and Hannes Ovrén. 2016. Efficient multi-frequency phase unwrapping using kernel density estimation. In *European Conference on Computer Vision.* Springer, 170–185.

Achuta Kadambi, Jamie Schiel, and Ramesh Raskar. 2016. Macroscopic interferometry: Rethinking depth estimation with frequency-domain time-of-flight. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* 893–902.

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 Transactions on Graphics (ToG)* 32, 6 (2013), 1–10.

Robert Lange. 2000. 3D time-of-flight distance measurement with custom solid-state image sensors in CMOS/CCD-technology. (2000).

Fengqiang Li, Huaijin Chen, Adithya Pediredla, Chiakai Yeh, Kuan He, Ashok Veeraraghavan, and Oliver Cossairt. 2017. CS-ToF: High-resolution compressive time-of-flight imaging. *Optics express* 25, 25 (2017), 31096–31110.

Fengqiang Li, Florian Willomitzer, Prasanna Rangarajan, Mohit Gupta, Andreas Velten, and Oliver Cossairt. 2018a. SH-ToF: Micro resolution time-of-flight imaging with superheterodyne interferometry. In *2018 IEEE International Conference on Computational Photography (ICCP).* 1–10. doi:10.1109/ICCPHOT.2018.8368473

Larry Li et al. 2014. Time-of-flight camera–an introduction. *Technical white paper* (2014), 10.

Yanlu Li, Jinghao Zhu, Matthieu Duperron, Peter O'Brien, Ralf Schüler, Soren Aasmul, Mirko De Melis, Mathias Kersemans, and Roel Baets. 2018b. Six-beam homodyne laser Doppler vibrometry based on silicon photonics technology. *Optics Express* 26, 3 (2018), 3638–3645.

Tzu-Tao Lin, Tzu-Hsien Sang, Chia-Ming Tsai, Gray Lin, and Sheng-Di Lin. 2025. SPAD ToF Array for High-Performance Non-contact Vibration Sensing. In *2025 IEEE International Electron Devices Meeting (IEDM)*. IEEE, 1–4.

Xiaomeng Liu, Kristofer Henderson, Joshua Rego, Suren Jayasuriya, and Sanjeev Koppal. 2021. Dense lissajous sampling and interpolation for dynamic light-transport. *Optics Express* 29, 12 (2021), 18362–18381.

Richard Lyons. 2008. Quadrature signals: complex, but not complicated. *URL: http://www. dspguru. com/info/tutor/quadsig. htm* (2008).

Peter Mahnke. 2018. Characterization of a commercial software defined radio as high frequency lock-in amplifier for FM spectroscopy. *Review of Scientific Instruments* 89, 1 (2018).

Radek Martinek, Lukas Danys, and Rene Jaros. 2019. Visible light communication system based on software defined radio: Performance study of intelligent transportation and indoor applications. *Electronics* 8, 4 (2019), 433.

Anthony F Martone, R Michael Buehrer, and David M McNamara. 2025. Emerging trends in radar: Metacognitive radar networks for the next generation of intelligent sensing. *IEEE Aerospace and Electronic Systems Magazine* (2025).

Parsa Mirdehghan, Brandon Buscaino, Maxx Wu, Doug Charlton, Mohammad E Mousa-Pasandi, Kiriakos N Kutulakos, and David B Lindell. 2024. Coherent optical modems for full-wavefield lidar. In *SIGGRAPH Asia 2024 Conference Papers*. 1–10.

Joseph Mitola. 2002. The software radio architecture. *IEEE Communications magazine* 33, 5 (2002), 26–38.

Nikhil Naik, Achuta Kadambi, Christoph Rhemann, Shahram Izadi, Ramesh Raskar, and Sing Bing Kang. 2015. A light transport model for mitigating multipath interference in time-of-flight sensors. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 73–81.

Desmond O'Connor. 2012. *Time-correlated single photon counting*. Academic press.

Ouster, Inc. 2019. *VLP-16 User Manual*. Ouster, Inc. https://data.ouster.io/downloads/velodyne/user-manual/vlp-16-user-manual-revf.pdf Document 63-9243 Rev. F (Accessed: 2026-08-21).

Rogério Passy, Nicholas Gisin, Jean-Pierre von der Weid, and HH Gilgen. 1994. Experimental and theoretical investigations of coherent OFDR with semiconductor laser sources. *Journal of lightwave technology* 12, 9 (1994), 1622–1630.

Adithya K Pediredla, Aswin C Sankaranarayanan, Mauro Buttafava, Alberto Tosi, and Ashok Veeraraghavan. 2018. Signal processing based pile-up compensation for gated single-photon avalanche diodes. *arXiv preprint arXiv:1806.07437* (2018).

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

Bhaskara Ramchander Rao. 2025. *Real-Time Signal Processing and State Estimation for Spaceflight Applications*. Ph.D. dissertation. Texas A&M University. https://hdl.handle.net/1969.1/1599617.

Christian Rembe, Benjamin J Halkon, and Mohamed AA Ismail. 2025. Measuring Vibrations in Large Structures with Laser-Doppler Vibrometry and Unmanned Aerial Systems: A Review and Outlook. *Advanced Devices & Instrumentation* 6 (2025), 0103.

Antony W Rix, John G Beerends, Michael P Hollier, and Andries P Hekstra. 2001. Perceptual evaluation of speech quality (PESQ)-a new method for speech quality assessment of telephone networks and codecs. In *2001 IEEE international conference on acoustics, speech, and signal processing. Proceedings (Cat. No. 01CH37221)*, Vol. 2. IEEE, 749–752.

Carl W Rossler, Emre Ertin, and Randolph L Moses. 2011. A software defined radar system for joint communication and sensing. In *2011 IEEE RadarCon (RADAR)*. IEEE, 1050–1055.

Steve J Rothberg, MS Allen, Paolo Castellini, Dario Di Maio, JJJ Dirckx, DJ Ewins, Ben J Halkon, P Muyshondt, Nicola Paone, T Ryan, et al. 2017. An international review of laser Doppler vibrometry: Making light work of vibration measurement. *Optics and Lasers in Engineering* 99 (2017), 11–22.

Daniel A Russell. 2000. On the sound field radiated by a tuning fork. *American Journal of Physics* 68, 12 (2000), 1139–1145.

Nofre Sanmartin-Vich, Javier Calpe, and Filiberto Pla. 2024. Analyzing the effect of shot noise in indirect Time-of-Flight cameras. *Signal Processing: Image Communication* 122 (2024), 117089. doi:10.1016/j.image.2023.117089

Tim Schenk. 2008. *RF imperfections in high-rate wireless systems: impact and digital compensation*. Springer.

Mark Sheinin, Dorian Chan, Matthew O'Toole, and Srinivasa G Narasimhan. 2022. Dual-shutter optical vibration sensing. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 16324–16333.

Shikhar Shrestha, Felix Heide, Wolfgang Heidrich, and Gordon Wetzstein. 2016. Computational imaging with multi-camera time-of-flight systems. *ACM Transactions on Graphics (ToG)* 35, 4 (2016), 1–11.

Julius Orion Smith. 2007. *Introduction to digital filters: with audio applications*. Vol. 2. Julius Smith.

AB Stanbridge and DJ Ewins. 1999. Modal testing using a scanning laser Doppler vibrometer. *Mechanical systems and signal processing* 13, 2 (1999), 255–270.


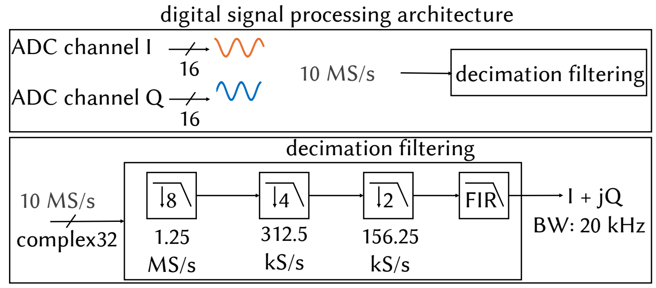


Fig. 13. Digital signal processing architecture. Complex IQ samples acquired at 10 MS/s are processed through a multistage FIR decimation chain reducing the sampling rate to 156.25 kS/s. Together with a final FIR low-pass filter, we trade our measurement bandwidth (from 5 MHz to 20 kHz) for improved SNR. FIR: finite-impulse response low-pass filter.


Robert W Stewart, Kenneth W Barlee, and Dale SW Atkinson. 2015. *Software defined radio using MATLAB & Simulink and the RTL-SDR*. Strathclyde Academic Media.

Alireza Tarighat, Rahim Bagheri, and Ali H Sayed. 2005. Compensation schemes and performance analysis of IQ imbalances in OFDM receivers. *IEEE Transactions on Signal Processing* 53, 8 (2005), 3257–3268.

Texas Instruments. 2014. Introduction to the Time-of-Flight (ToF) System Design. *Texas Instruments User's Guide* SBAU219D (2014), 1–47.

Brevin Tilmon, Eakta Jain, Silvia Ferrari, and Sanjeev Koppal. 2020. Foveacam: A mems mirror-enabled foveating camera. In *2020 IEEE International Conference on Computational Photography (ICCP)*. IEEE, 1–11.

Walter HW Tuttlebee. 2002. *Software defined radio: enabling technologies*. John Wiley & Sons.

Rini Khamimatul Ula, Yusuke Noguchi, and Koichi Iiyama. 2019. Three-dimensional object profiling using highly accurate FMCW optical ranging system. *Journal of Lightwave Technology* 37, 15 (2019), 3826–3833.

Tore Ulversoy. 2010. Software defined radio: Challenges and opportunities. *IEEE Communications Surveys & Tutorials* 12, 4 (2010), 531–550.

Deepak Vasisht, Swarun Kumar, and Dina Katabi. 2016. {Decimeter-Level} localization with a single {WiFi} access point. In *13th USENIX symposium on networked systems design and implementation (NSDI 16)*. 165–178.

Neal Wadhwa, Michael Rubinstein, Frédo Durand, and William T Freeman. 2013. Phase-based video motion processing. *ACM Transactions on Graphics (ToG)* 32, 4 (2013), 1–10.

Neal Wadhwa, Michael Rubinstein, Frédo Durand, and William T Freeman. 2014. Riesz pyramids for fast phase-based video magnification. In *2014 IEEE International Conference on Computational Photography (ICCP)*. IEEE, 1–10.

Qing Wang, Domenico Giustiniano, and Daniele Puccinelli. 2014. OpenVLC: Software-defined visible light embedded networks. In *Proceedings of the 1st ACM MobiCom workshop on Visible light communication systems*. 15–20.

Marcus Windisch and Gerhard Fettweis. 2004. Blind I/Q imbalance parameter estimation and compensation in low-IF receivers. In *First International Symposium on Control, Communications and Signal Processing, 2004*. IEEE, 75–78.

Shengxi Wu, Sophia Yang, Dorian Chan, and Matthew O'Toole. 2026. Computational Speckle Pattern Interferometry. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*.

Fu-Min Zhang, Ya-Ting Li, Hao Pan, Chun-Zhao Shi, and Xing-Hua Qu. 2019. Vibration compensation of the frequency-scanning-interferometry-based absolute ranging system. *Applied Sciences* 9, 1 (2019), 147.

Tianyuan Zhang, Mark Sheinin, Dorian Chan, Mark Rau, Matthew O'Toole, and Srinivasa G Narasimhan. 2023. Analyzing physical impacts using transient surface wave imaging. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 4339–4348.

Xinyu Zhou, Peiqi Duan, Yeliduosi Xiaokaiti, Chao Xu, and Boxin Shi. 2025. Event-based Visual Vibrometry. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 24666–24676.


# A Digital signal processing pipeline

## A.1 Filter design

The USRP N210 provides dual-channel ADC measurements which we configure to stream 16-bit complex IQ samples at a sampling rate of 10 MS/s[7]. The physical information of interest is encoded in the slowly varying phase of this complex envelope, and therefore, the full input sampling rate is not required for subsequent processing.

We process the complex signal using a multistage causal decimation chain shown in Fig. 13. The discrete-time IQ signal sampled at 10 MS/s is passed through a sequence of anti-aliasing low-pass filters followed by downsampling stages with factors $[8, 4, 2]$, producing intermediate sampling rates of 1250, 312.5, and a final rate of 156.25 kS/s.

Each stage employs a Kaiser-window finite impulse response (FIR) low-pass filter designed to preserve the measurement bandwidth while suppressing spectral components that would alias after decimation. The filters retain a passband up to 25 kHz with a 5 kHz transition band to ensure that the measurement bandwidth of 20 kHz remains unaffected.

A multistage architecture significantly reduces computational complexity compared to single-stage decimation [Ramchander Rao 2025; Smith 2007]. Directly filtering the 10 MS/s stream to the final bandwidth would require extremely long FIR filters due to the narrow normalized transition width. By progressively reducing the sampling rate, each stage operates with a wider normalized transition band, enabling practical filter lengths and efficient implementation.

All filtering operations are implemented causally using FIR convolution followed by integer-rate downsampling, making the architecture suitable for real-time streaming and embedded deployment. After the final decimation stage, an additional FIR filter at 156.25 kS/s enforces the desired measurement bandwidth prior to phase extraction.

Overall, the chain reduces the sample rate by a factor of 64 while preserving the phase dynamics within the 20 kHz measurement bandwidth. The final sampling rate of 156.25 kS/s corresponds to approximately 7.8 samples per cycle at the highest measurement frequency, providing sufficient temporal resolution for stable phase tracking and reconstruction of audio-band phase variations.

After this stage removes high-frequency noise from the IQ samples, the remaining phase signal primarily exhibits low-frequency drift introduced by oscillator mismatch in the SDR demodulation chain. We address this drift using the compensation method described next.

## A.2 Digital phase-locked loop

The phase recovered from coherent demodulation contains the displacement information of target. However, in practice, the measured phase exhibits a slowly varying drift. This drift arises from residual frequency mismatch between the transmit and receive oscillators as well as oscillator phase noise.

Let the recovered phase be

$$\phi[n] = \phi_{\text{motion}}[n] + \phi_{\text{drift}}[n] \tag{24}$$

where $\phi_{\text{motion}}[n]$ represents the phase dynamics caused by displacement and $\phi_{\text{drift}}[n]$ is a slow varying phase trend due to oscillator mismatch.

Although the nominal frequency accuracy of the commercial SDRs are on the order of a few parts per million, even small offsets produce significant phase drift over time. For example, a frequency offset $\Delta f$ produces a phase ramp of $\phi_{\text{drift}}[n] = 2\phi\Delta f n T_s$ ($n$ - sample index and $T_s$ is sampling period).

In practice, the observed drift is not strictly linear. Temperature variations, oscillator instability, and phase noise introduce *nonlinear low-frequency phase fluctuations* that cannot be accurately removed by fitting and subtracting a linear ramp.

Consequently, a more robust approach is required to continuously estimate and remove this slow phase evolution while preserving higher-frequency phase variations corresponding to the physical motion.

To accomplish this, we employ a digital phase-locked loop (PLL).

*A.2.1 PLL architecture.* A PLL is a closed-loop feedback control system that continuously adjusts the phase of a locally generated oscillator to match the phase of an incoming signal [Gardner 2005]. In this work, the PLL is configured with a very low loop bandwidth so that it tracks only the slow oscillator drift while leaving higher-frequency phase components unchanged. The block design is outlined in Fig. 14.

*A.2.2 Signal model.* Described previously in Fig. 2, the SDR receiver produces digital samples of the complex baseband signal ($f_b$). Consider the received complex baseband signal be $s_b[n]$

$$s_b[n] = a_{\text{tx}} e^{j(2\pi f_b n T_s + \phi[n])}. \tag{25}$$

The goal of the PLL is to estimate the slowly varying component of $\phi[n]$ which contains both motion information and drift.

*A.2.3 Numerically controlled oscillator (NCO)..* The NCO generates a complex reference signal

$$l[n] = e^{-j(2\pi f_b n T_s + \hat{\phi}[n])}, \tag{26}$$

where $\hat{\phi}[n]$ is the phase estimate produced by the PLL.

The NCO is essentially a digital oscillator whose phase can be continuously adjusted by the feedback loop. By modifying $\hat{\phi}[n]$, the PLL can align the local oscillator phase with the incoming signal.

*A.2.4 Digital downconversion.* The received signal is multiplied by the NCO output:

$$\tilde{s}_b[n] = s_b[n] l[n]. \tag{27}$$

Substituting the expression from Eq. (26) in Eq. (27),

$$\tilde{s}_b[n] = e^{j(\phi[n] - \hat{\phi}[n])}, \tag{28}$$

the output phase becomes

$$\Delta\phi[n] = \phi[n] - \hat{\phi}[n]. \tag{29}$$

If the PLL successfully tracks the drift component, $\hat{\phi}[n] \approx \phi_{\text{drift}}[n]$, and the residual phase $\Delta\phi[n]$ contains only the motion dynamics.

*A.2.5 Phase detector.* The phase detector estimates the phase difference between the incoming samples and the samples locally produced by the NCO. For a complex downconverted signal $\tilde{s}_b[n]$, the phase error $e[n]$ is modeled using the normalized quadrature component

$$e[n] = \text{Imag}\Big(\frac{\tilde{s}_b[n]}{|\tilde{s}_b[n]|}\Big), \tag{30}$$

[7] ADC precision is 14-bits, USRP packets 16-bit wide.

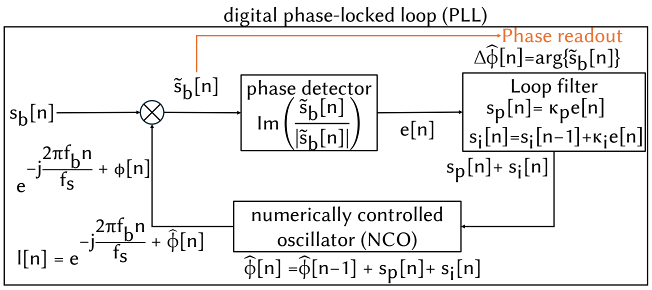


Fig. 14. **Digital phase-locked loop (PLL) used for drift compensation**. The received baseband signal $s_b[n]$ is mixed with a numerically controlled oscillator (NCO) to obtain the downconverted signal $\tilde{s}_b[n]$. A quadrature phase detector computes the phase error $e[n]$, which is filtered by a proportional–integral (PI) loop filter to update the NCO phase estimate $\hat{\phi}[n]$. The loop tracks slow oscillator drift while preserving higher-frequency phase variations corresponding to physical motion, which are obtained from the residual phase $\Delta\hat{\phi}[n] = \arg\{\tilde{s}_b[n]\}$. The results are captured in Fig. 5 of the paper.

which turns out to be $e[n] = \sin(\Delta\phi[n])$ and for small phase errors, we know that $\sin(\Delta\phi) \approx \Delta\phi$.

So the phase detector output provides a measure of the instantaneous phase error between the incoming signal and the NCO. This phase error forms the input to the feedback controller.

*A.2.6 Loop filter (PI controller).* The loop filter provides control actuation in correcting the phase and frequency differences to lock on to the incoming samples. The filter performs two functions: (1) the proportional term corrects instantaneous phase differences, (2) the integral term accumulates error to compensate frequency offsets.

The loop update equations are

$$s_p[n] = \kappa_p e[n] \qquad \text{(proportional action)}, \tag{31}$$

$$s_i[n] = s_i[n-1] + \kappa_i e[n] \qquad \text{(integral action)}, \tag{32}$$

$$\hat{\phi}[n] = \hat{\phi}[n-1] + s_p[n] + s_i[n], \qquad \text{(phase correction)} \tag{33}$$

where $\kappa_p, \kappa_i$ are the proportional and integral gains, respectively. $s_p[n], s_i[n]$ denote the signals from the proportional and integral sections of the loop filter, respectively. We select the loop filter parameters to achieve a desired loop bandwidth and damping factor, which together determine the dynamic response of the PLL.

*A.2.7 Loop bandwidth.* The loop bandwidth ($B_L$) determines which phase variations are tracked by the PLL: components below $B_L$ are tracked by the loop while components above $B_L$ remain in the residual phase output.

The digital PLL implemented in this work follows a second-order type-II control loop. Gardner [Gardner 2005] covers a detailed overview of digital PLLs. The continuous-time closed-loop transfer function of a second-order PLL, in Laplace domain, can be written as

$$H(s) = \frac{2\zeta\omega_n s + \omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}, \tag{34}$$

where $\omega_n$, the natural frequency of the loop, determines how quickly the loop can correct phase and frequency errors, $\zeta$ is the damping factor that determines the transient behavior of the loop, governing the loop behavior in correcting phase errors. In practice, PLLs are commonly designed with $\zeta \approx 0.707$ to achieve a stable and well-damped response.

As shown in [Stewart et al. 2015], for a digital PLL operating at sampling frequency $f_s$, the loop parameters are expressed in terms of the normalized bandwidth $\theta = B_L/f_s$.

The proportional and integral gains of the loop filter are then obtained from the standard discretization of the second-order loop dynamics. The gains used in the implementation (Eqs. (31) to (33)) are

$$\kappa_p = \frac{4\zeta\theta}{1 + 2\zeta\theta + \theta^2}, \tag{35}$$

$$\kappa_i = \frac{4\theta^2}{1 + 2\zeta\theta + \theta^2}. \tag{36}$$

*A.2.8 Phase readout.* After PLL locks onto the phase and frequency of the incoming signals, the residual phase is

$$\Delta\phi[n] = \phi[n] - \hat{\phi}[n], \tag{37}$$

which corresponds to the drift compensated phase.

We compute the final phase readout as $\phi_{\text{out}}[n] = \arg\{y[n]\}$. This signal contains the desired displacement-induced phase variations with the slow oscillator drift removed.