

Snapshot Lidar: Fourier embedding of amplitude and phase for single-image depth reconstruction

Sarah Friday^{1*}, Yunzi Shi^{1*}, Yaswanth Cherivirala², Vishwanath Saragadam³, Adithya Pediredla¹

*equal contribution, ¹Dartmouth College, ²University of Michigan, Ann Arbor, ³University of California, Riverside

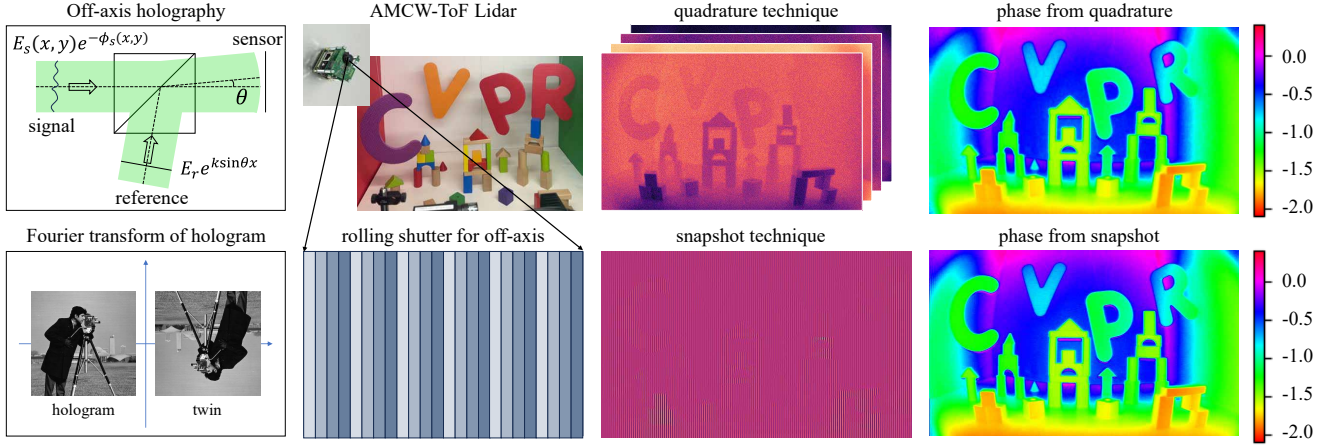


Figure 1. We propose a novel snapshot Lidar imaging system that is inspired by off-axis holography techniques. (1st col.) Off-axis holography uses oblique illumination to separate the hologram and its twin in Fourier space. (2nd col.) We leverage the rolling-shutter effect of amplitude modulated continuous wave time-of-flight (AMCW-ToF) cameras to emulate the off-axis principle, thereby separating the time-of-flight hologram and its twin in Fourier space. (3rd col.) The conventional operation of AMCW-ToF Lidars requires four measurements, whereas our technique is 4x faster, improving both the data bandwidth and temporal resolution. (4th col.) Reconstructed phase (that encodes depth) from measurements—The reconstructions are quite similar even with 4x fewer measurements.

Abstract

Amplitude modulated continuous-wave time-of-flight (AMCW-ToF) cameras are finding applications as flash Lidars in autonomous navigation, robotics, and AR/VR applications. A conventional CW-ToF camera requires illuminating the scene with a temporally varying light source and demodulating a set of quadrature measurements to recover the scene’s depth and intensity. Capturing the four measurements in sequence renders the system slow, invariably causing inaccuracies in depth estimates due to motion in the scene or the camera. To mitigate this problem, we propose a snapshot Lidar that captures amplitude and phase simultaneously as a single time-of-flight hologram. Uniquely, our approach requires minimal changes to existing CW-ToF imaging hardware. To demonstrate the efficacy of the proposed system, we design and build a lab prototype, and evaluate it under varying scene geometries, illumina-

tion conditions, and compare the reconstructed depth measurements against conventional techniques. We rigorously evaluate the robustness of our system on diverse real-world scenes to show that our technique results in a significant reduction in data bandwidth with minimal loss in reconstruction accuracy. As high-resolution CW-ToF cameras are becoming ubiquitous, increasing their temporal resolution by four times enables robust real-time capture of geometries of dynamic scenes.

1. Introduction

Amplitude modulated continuous-wave time-of-flight (AMCW-ToF) cameras, also known as correlation-based time-of-flight or indirect time-of-flight cameras, are used as flash Lidars to compute scene depth, and are used in autonomous navigation, robotics [6, 13], and augmented reality [8]. AMCW-ToF cameras operate by projecting a

	illumination	reference	physics
holography	coherent beam	planar beam	optical interference
off-axis	coherent beam	tilted beam	optical interference
CW-ToF	AMCW	electronic shutter	electronic multiplication
proposed	AMCW	rolling shutter	electronic multiplication

Table 1. **Principles of holography, off-axis holography, CW-ToF, and proposed snapshot CW-ToF.** We draw parallels between off-axis holographic setup, and our proposed Fourier embedded ToF, that allow us to show that using a rolling-shutter sensor and varying the reference phase during CW-ToF acquisition results in a snapshot capture of the “ToF hologram” that contains both the amplitude and depth (encoded in phase) of the scene.

temporally varying (often a sinusoidal) light source, and then correlating it on the sensor side with an appropriate (also often a sinusoid) decoding function. Depth is encoded in the phase of the measurements, and hence up to four measurements (quadrature) are required to robustly estimate the depth and intensity of the scene. These quadrature measurements are captured by temporal multiplexing, which invariably leads to lower frame rates and suffers from motion artifacts.

Inspired by snapshot off-axis holography techniques, we propose a snapshot CW-ToF imaging technique that measures amplitude and depth using a single capture. The key observation is that by defining the *time-of-flight hologram* as the complex sinusoid whose amplitude and phase are proportional to the intensity and depth of the scene, we can draw parallels between holography and CW-ToF imaging techniques. This novel formulation of the ToF hologram enables us to translate off-axis techniques to CW-ToF imaging. In particular, by using rolling-shutter CW-ToF sensors [1] and spatially varying the reference phase of the coded exposure, we can emulate the off-axis holography effect and recover the complex-valued ToF hologram. In [Tab. 1](#), we highlight the similarities and differences between holography, off-axis holography, CW-ToF, and the proposed imaging techniques.

To achieve snapshot depth imaging, we leverage the image formation model of CW-ToF cameras, and demonstrate our modification to enable the capture of a ToF hologram with unipolar (all-positive decoding function) and bipolar (positive and negative decoding function) codes. We then show analytical and computational techniques to recover amplitude and depth from a rolling-shutter image and the need for optical prefiltering to prevent aliasing and noise folding.

We design and build a hardware setup built in the lab

with a Melexis 75027 device with region-of-interest support and a galvanometer to emulate the rolling-shutter effect. Using this device, in [Sec. 6](#), we demonstrate that the proposed technique reduces data bandwidth and improves frame rate on various scenes containing diffuse, specular, and refractive objects. The proposed technique is robust to dead and saturated pixels, thereby enabling depth imaging with slightly faulty sensors, and extremely bright settings. We carefully evaluate design parameters including prefiltering window size, and spatial phase variation rate, and provide an experimentally optimal set of values that are agnostic to scene conditions. We empirically show that the proposed Fourier-based reconstruction technique is superior to the standard N-bucket technique for reconstruction.

2. Related Work

Continuous wave time-of-flight cameras. CW-ToF cameras measure depth at each spatial pixel with a temporally modulated light source. The intensity of the scene is encoded in the amplitude, and depth in the phase, of the measurements. Typically four or more images, with different phases, are required to recover both amplitude and depth. These four measurements are obtained either with a spatially multiplexed sensor (similar to a Bayer pattern), or with sequential measurements [12]. Spatial multiplexing results in severe aliasing artifacts, and is inherently expensive and cumbersome to manufacture. In contrast, sequential measurements invariably result in motion artifacts when capturing dynamic scenes.

Advances in CWToF sensing With the increasing number of practical applications for CWToF, work has been done to enhance its capabilities and mitigate its limitations. CS-ToF [20] is a compressive ToF imaging architecture aimed at overcoming sensor resolution limitations. As compressive sensing relies on the assumption of a linear measurement process for high resolution image estimation, CS-ToF uses a phasor representation inspired by Gupta [9] and O’Toole [25] of the ToF output to create a linear model between the scene and ToF measurements. In a process similar to single pixel imaging, laser light is reflected off of an object onto a high resolution digital micro-mirror device (DMD), and then relayed onto a lower resolution ToF sensor. By changing the DMD codes across multiple exposures, CS-ToF performs spatiotemporal multiplexing of the scene’s amplitude and phase; however, sacrificing temporal resolution for spatial resolution. CWToF has also been combined with other modalities such as spectrum [27], light transport [24, 31], and light fields [15] that have significantly expanded the applications of CWToF cameras. Nevertheless, it is worth noting that there are few, if any, approaches that capture CWToF data in a snapshot manner, that is crucial for dynamic scenes.

Superheterodyned interferometry. A related technology that enables snapshot ToF sensing is superheterodyned interferometric time-of-flight (SH-ToF) [21, 22], where two closely spaced illumination wavelengths combine to produce a much larger (heterodyned) synthetic wavelength. This approach however still requires capturing two separate images (one with each wavelength) to produce the final superheterodyned measurement. Recent work by Ballester et al. [3] demonstrated that SH-ToF can be combined with the off-axis holography to obtain a snapshot ToF imaging system. We next briefly discuss off-axis holography to understand its relevance to ToF sensing.

Off-axis holography Off-axis holography [7, 19] is a widely known imaging technique for reconstructing the amplitude ($E_s(x, y)$) and phase ($\phi_s(x, y)$) of a hologram with a single measurement. The experimental setup schematic is shown in the first column of Fig. 1. The measurement by the camera is given by

$$\begin{aligned} I &= |E_s(x, y)e^{-\phi_s(x, y)} + E_r e^{-k \sin(\theta)x}|^2 \\ &= |E_s(x, y)|^2 + |E_r|^2 + E_r^* E_s(x, y) e^{k \sin(\theta)x - \phi_s(x, y)} \\ &\quad + E_r E_s^*(x, y) e^{-k \sin(\theta)x + \phi_s(x, y)} \end{aligned} \quad (1)$$

Off-axis holography embeds the hologram ($E_s(x, y)e^{-\phi_s(x, y)}$) and its twin ($E_s^*(x, y)e^{\phi_s(x, y)}$) separately in the Fourier domain allowing us to recover the hologram computationally. Ballester et al. [2] showed that off-axis techniques can be used in synthetic wavelength interferometry. Specifically, the system consisted of a colocated, dual wavelength illumination source, and two, spatially separated reference beams. The separated location implied that the phasor information was encoded in different parts of the Fourier spectrum, enabling a simple frequency domain post-processing algorithm to estimate depth information with a single image. Our technique is inspired by off-axis holography. Two key observations that enable our technique are (i) expressing the time-of-flight image formation model with complex sinusoids allows us to draw parallels between holography and AMCW-ToF (ii) the rolling-shutter effect allows us to emulate off-axis technique.

Applications of CW-ToF cameras. Beyond depth imaging, CW-ToF cameras enable a plethora of important applications in imaging. For example, Callenberg et al. [4] showed that a CW-ToF camera can enable difference imaging, that is often useful for implementing convolutional operations within sensors. Achar et al. [1] showed that ToF sensors can be used to tease out sub-surface features through epipolar gating. Similarly, Naik et al. [24] and Kadambi et al. [16] showed that a light transport matrix-based formulation of ToF measurements enabled measur-

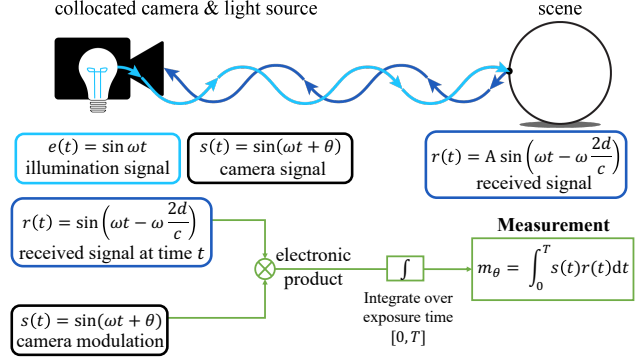


Figure 2. **CW-ToF imaging systems.** CW-ToF cameras use high-frequency illumination sources and exposure codes to indirectly measure the object’s depth. The measurement is the product of the delayed illumination signal and programmatically phase-shifted exposure code. The delay encodes object depth and is recoverable by taking multiple measurements with varying phase-shift (θ).

ing multipath interferences, that are key to measuring geometric properties of complex objects like metal and glass. Our proposed work enables a similar slew of applications beyond depth imaging, including sub-surface imaging and multipath interference reduction (via epipolar gating).

3. Technical Background

As shown in Fig. 2, an amplitude-modulated continuous-wave time-of-flight camera (AMCW-ToF) has an illumination source whose amplitude ($e(t)$) varies over time at several tens of MHz. While these amplitude variations are typically sinusoidal, other modulations have been successfully demonstrated in the past [9–11] for various applications. The light received ($r(t)$) at the sensor is attenuated and delayed based on the light transport of the scene ($r(t) = Ae(t - \tau)$), where τ is the total time-of-travel of the light beam. Assuming a colocated light source and sensor and no multi-bounce paths, $\tau = 2d\omega/c$, where d is the distance to the object, ω is the modulation frequency, and c is the speed of light. The sensor temporal exposure ($s(t)$) is also modulated, typically as a square wave (unipolar or bipolar) with the same frequency as the illumination. Recent sensors mostly use bipolar coding with a two-bucket architecture [30], and the exposure duration is approximated with a sinusoidal modulation. The measurement by the sensor is given by

$$\begin{aligned} m_\theta &= \frac{1}{T} \int_{t=0}^T Ar(t)s(t)dt \\ &= \frac{A}{T} \int_{t=0}^T \sin(\omega t + \phi) \sin(\omega t + \theta)dt = \frac{A}{2} \cos(\theta - \phi). \end{aligned} \quad (2)$$

Most hardware implementations of CWToF Lidar mea-

sure four images m_0 , $m_{\pi/2}$, m_π , and $m_{3\pi/2}$ and compute amplitude and phase as

$$A = 2\sqrt{(m_0 - m_\pi)^2 + (m_{\pi/2} - m_{3\pi/2})^2}$$

$$\phi = \arctan \frac{m_{\pi/2} - m_{3\pi/2}}{m_0 - m_\pi}. \quad (3)$$

The quadrature technique is also extended to the N-tap technique [28] by taking N measurements corresponding to $\theta = \{0, 2\pi/N, 4\pi/N, \dots, 2\pi(N-1)/N\}$ and computing amplitude and phase as:

$$A = \sqrt{\left(\sum_{n=0}^{N-1} m_{\frac{2\pi n}{N}} \cos \frac{2\pi n}{N}\right)^2 + \left(\sum_{n=0}^{N-1} m_{\frac{2\pi n}{N}} \sin \frac{2\pi n}{N}\right)^2}$$

$$\phi = \arctan \frac{\sum_{n=0}^{N-1} m_{\frac{2\pi n}{N}} \cos \frac{2\pi n}{N}}{\sum_{n=0}^{N-1} m_{\frac{2\pi n}{N}} \sin \frac{2\pi n}{N}} \quad (4)$$

Unlike these techniques, we reconstruct amplitude and phase using a single image, which we discuss next.

4. Proposed Technique: Snapshot CWToF

The key idea for measuring amplitude and depth using a single image is to spatially vary the phase θ linearly along rows or columns during the exposure. We can achieve this spatial variation with simple hardware modifications by either (1) changing the existing hardware design with different exposure phase offsets per row/column, (2) using a rolling-shutter camera and changing the phase offset of either the illumination or sensor during each row capture, or (3) using a fast line sensor (or hardware region-of-interest supporting camera) and scanning the scene line by line with a Galvanic mirror while changing the phase offset of each line. We used the last alternative in this paper, as the Melexis 75027 supports hardware ROI [23].

To understand how linearly varying θ embeds both amplitude and phase in the captured image, we follow the notation from O'Toole et al. [26], Li et al. [21], and Gupta et al. [9] and represent the amplitude and phase images with complex sinusoidal notation as $I(x, y) = A(x, y)e^{-j2d\omega/c}$ and the Fourier Transform of this sinusoid as $\mathcal{I}(\omega_x, \omega_y)$. For brevity, we define $\phi = 2d\omega/c$ as the phase shift of the illumination signal. Varying θ linearly during the measurement results in

$$m_{\theta=kx}(x, y) = \frac{A(x, y)}{2} \cos(kx - \phi(x, y))$$

$$= \frac{A(x, y)}{4} \left(e^{j(kx - \phi(x, y))} + e^{-j(kx - \phi(x, y))} \right) \quad (5)$$

Taking the Fourier transform on both sides, we get

$$\mathcal{F}_{m_{kx}}(\omega_x, \omega_y) = \frac{1}{4} (\mathcal{I}(\omega_x - k, \omega_y) + \mathcal{I}^*(k - \omega_x, -\omega_y)), \quad (6)$$

which is similar to the off-axis holography Eq. (1). Therefore, we will refer to the complex sinusoidal notation ($I(x, y)$) as the *ToF hologram* and its complex conjugate ($I^*(x, y)$) as the *ToF twin*.

As depicted in Fig. 3, to recover the ToF hologram from the image captured by varying θ linearly, we take the Fourier transform of the measured image, filter the ToF twin, right shift the ToF hologram, and compute the inverse Fourier transform. The amplitude and phase of the inverse Fourier transform result are the amplitude and phase of the scene.

Extension to unipolar codes. Most recent ToF sensors use a two-bucket architecture to suppress background illumination; hence, the codes are bipolar. However, our technique is also amenable to unipolar-coded time-of-flight cameras and imperfect two-bucket architecture cameras. In this case, the measurement model becomes

$$m_\theta = \frac{A}{2} \cos(\theta - \phi) + a. \quad (7)$$

where a is proportional to the ambient background light intensity. In this case, the Fourier transform has a DC component that can be filtered out along with the ToF twin.

Aliasing, noise folding, and prefiltering. From Fig. 3 (b), we notice that the ToF hologram and twin overlap in the Fourier domain. This overlap results in aliasing, where the high-frequency content of the ToF twin may appear as low-frequency content of the ToF hologram, and noise folding, where the twin's noise folds into the ToF hologram's noise.

To mitigate aliasing and noise folding, we optically filter the high-resolution ToF hologram before the linear phase shifting and measurement. In the hardware implementation, we achieve this by simply defocusing the imaging lens in front of the sensor. This simple blurring operation reduces aliasing and noise folding artifacts (Fig. 3 (d, e)).

Therefore, while the proposed snapshot Lidar technique can capture amplitude and phase with a single image, it does so by sacrificing spatial resolution for temporal resolution. As spatial resolutions of imaging sensors keep increasing (compare PMD vs DME660 vs. Melexis) due to better fabrication techniques, trading spatial resolution for improving temporal resolution is arguably more useful.

Phase noise. As with any optical measurements, the recovered phase will have noise due to shot noise in the measurements. This shot noise is present even for the standard N-bucket technique [17, 18]. The standard deviation of the phase noise for the snapshot technique is

$$\sigma_\phi^2 \approx \frac{\partial \phi}{\partial m_{kx}} m_{kx} \quad (8)$$

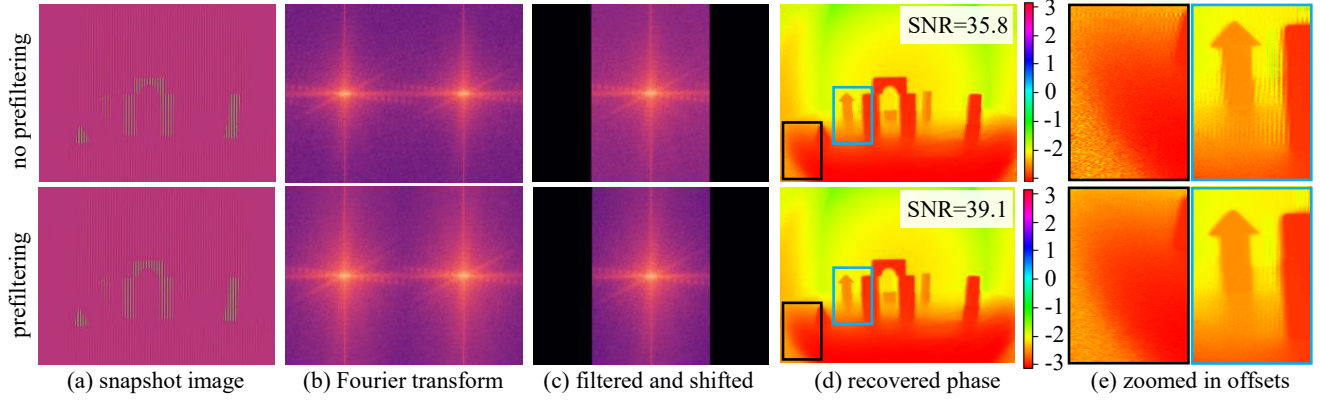


Figure 3. **Snapshot CWToF decoding, and effect of prefiltering.** (a) captured snapshot CW-ToF image. (b) Fourier transform of snapshot image. (c) We filter out the twin and frequency shift the hologram in the Fourier domain to recover the ToF hologram. (d) We reconstruct the phase by computing the phase of the inverse Fourier transform of (c). (top row vs. bottom row) We prefilter the ToF hologram using a defocused imaging lens, which can decrease the overlap between the hologram and its twin. The prefiltering decreases aliasing and noise folding, resulting in a 3dB SNR gain ($2\times$ smaller phase error). (e) zoomed-in offsets. (left) Without prefiltering, the high-frequency phase noise of the twin shows up as the low-frequency phase noise. (right) aliasing decreases with prefiltering.

where ϕ is explicitly expressed as

$$\phi(x, y) = \arctan \left(\frac{m_{kx}(x, y) \sin(kx) \otimes \text{sinc}(kx)}{m_{kx}(x, y) \cos(kx) \otimes \text{sinc}(kx)} \right).$$

Here, \otimes is spatial convolution. We show the derivation in the supplementary material. An analytical comparison of the noise variance between the N-bucket and snapshot techniques is not feasible. Instead, in Sec. 6, we show empirical comparison and show that the snapshot phase measurement does not result in additional phase noise.

5. Experimental setup

To implement our method, we use the EVK75027-110-940-2 evaluation kit (EVK) by Melexis which uses diffused 940nm VCSEL diodes for illumination and a Melexis 75027 CWToF sensor. While the sensor is a global shutter, we can create the rolling shutter effect by taking advantage of the sensor’s hardware programmable ROI to reduce the per-frame capture/readout to a single row.

For our experiments, we mount the EVK75027 to a linear 3D stage and carefully align the view to a set of 2-axis galvanic mirrors (GVS012) from Thorlabs. Then, we reduce the ROI of the camera to a single line, and steer the view of the galvos with a DAQ. Additionally, we separated the illumination board from the rest of the EVK and mounted it on top of the galvanic mirrors for approximate collocation of light source and apparent camera location. The MLX75027 module contained a firmware lock on the frame rate at 100 fps, severely limiting our capture speed when the ROI is reduced to only a single line. However, our setup demonstrates a proof-of-concept, and future versions of Melexis cameras without a firmware lock will enable a real-time operation.

One way to emulate snapshot measurements from our setup is to capture multiple images with varying phases and appropriately select the rows/columns in the image. In practice, we found it to be significantly simpler and convenient to conduct most of our experiments. In Fig. 4, we show the hardware prototype with fast-scanning galvanic mirrors. Additionally in Fig. 5, we compare results from an emulated versus hardware implemented snapshot capture and notice that the techniques result in the same amplitude and phase estimates. For more details about the hardware prototype and emulation method, we refer the reader to the supplemental material.

6. Experiments

In this section, we use the hardware setup and emulation technique described in Sec. 5 to evaluate the proposed snapshot technique systematically. We show that (1) prefiltering improves signal-to-noise ratio (SNR) and prefiltering with cylindrical lenses is better than regular lenses; (2) the optimal phase variation rate is ($k = \frac{\pi}{2} \text{line}^{-1}$); (3) the Fourier technique is superior to the N-bucket technique for snapshot images; (4) the snapshot technique does not add any additional noise and (5) is robust to saturated and dead pixels; (6) rotating the phase variation (by rotating the camera itself) decreases aliasing. In the supplementary, we show extended results and a comparison with the dual phase mode of the epc660 camera. We share our code on [GitHub](#).

Optimal prefiltering kernel. As mentioned in Sec. 4, prefiltering reduces aliasing and noise folding artifacts. As aliasing and noise folding occur in the direction of the rolling shutter, blurring the ToF hologram only in the rolling shutter direction is better than isotropic blurring as shown in Fig. 6. We can implement hardware anisotropic blurring by

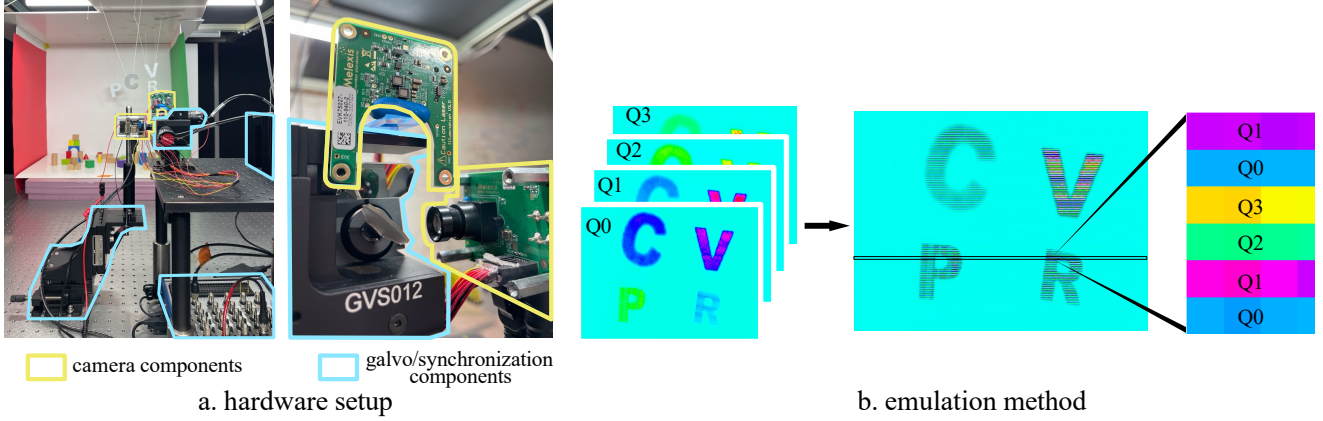


Figure 4. **Lab prototype, and rolling shutter encoding scheme.** (a) Rolling shutter hardware prototype with the camera components outlined in yellow and galvonic mirrors/synchronization hardware outlined in blue. (b) Compositing quadrature images to create an emulated rolling shutter snapshot.

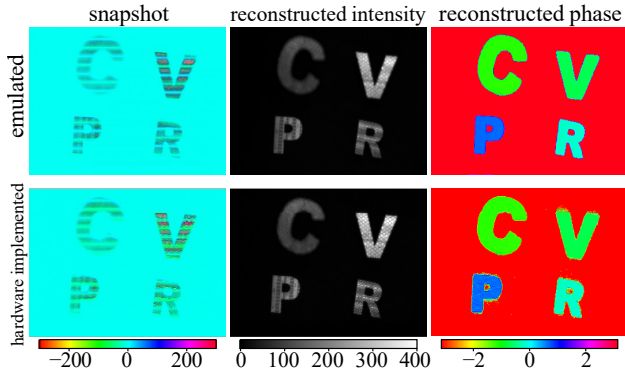


Figure 5. **Advantages of emulated snapshot capture.** Amplitude and depth reconstructed by snapshot data captured with emulations (this paper) vs. single line ROI (ideal optical setup) and galvanic mirrors. The emulated measurements and the reconstructed phase are similar to the full prototype. We hence demonstrated all experiments with an emulated setup, as it enables a faithful evaluation of a true snapshot system.

using a defocused cylindrical lens. Note that the optimal kernel for prefiltering is a 1D sinc function in the spatial domain (corresponding to eliminating all the overlapping frequencies). However, that is not physically possible. Designing the optimal optically-realizable kernel is beyond the scope of this paper.

Assuming a 1D-Gaussian prefiltering kernel, we computed the effect of blur kernel size (measured in standard deviation σ) on the phase reconstruction accuracy. As shown in Fig. 7 (a), for all values of k , the SNR initially improves when σ is increased and then decreases showing that an optimal blur kernel size exists. This phenomenon occurs as prefiltering initially reduces the overlap between the ToF hologram and the twin but later reduces the frequency con-

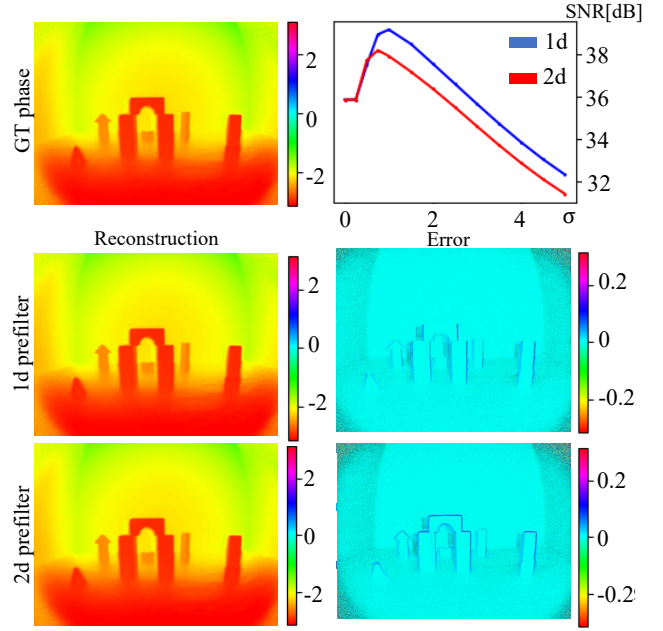


Figure 6. **Robustness to noise due to prefiltering.** Prefiltering reduces aliasing and noise folding artifacts to improve the reconstruction quality. Compared to the 2D-blur kernel, the 1D-blur kernel in the direction of rolling shutter direction results in a significantly higher-quality reconstruction. Since the ToF hologram and twin overlap occur only in the rolling shutter direction, a 1D blur kernel preserves the spatial frequencies in the direction orthogonal to rolling shutter.

tent in the captured ToF hologram.

Optimal phase variation rate k . In Fig. 7 (b), we plot optimal σ value as a function of $R = 2\pi/k$, which represents the average number of rows required before the θ

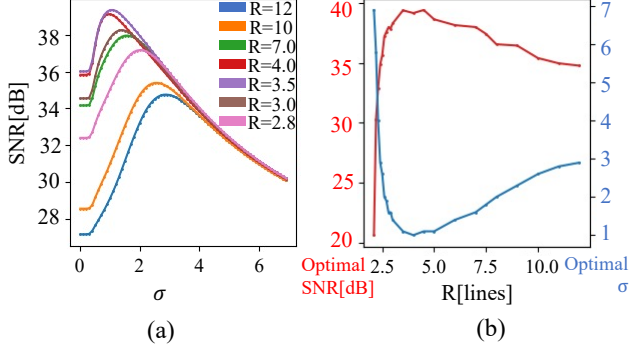


Figure 7. **Optimal prefilter size and phase variation rate.** (a) We show the effect of the standard deviation of the blur kernel (blur kernel size) on the phase reconstruction accuracy for various phase variation rates (k). The optimal prefilter size depends on the phase variation rate. (b) The best phase reconstruction rate that results in best reconstruction quality is around $R = 4$, and this is consistently observed for various scenes. $R = 4$ results in the least overlap between the ToF hologram and its twin.

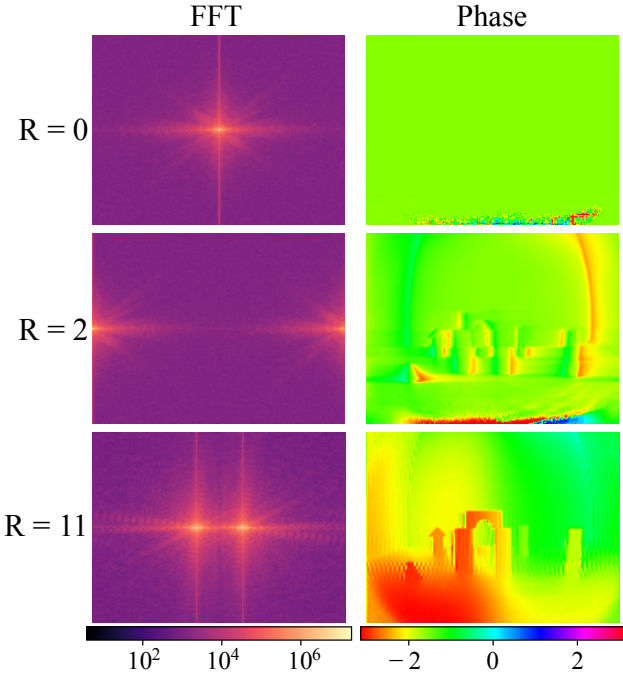


Figure 8. **Visualization of the effect of changing the phase variation rate.** In practice, we take the Discrete Fourier Transform which is circular. The smaller the overlap between the hologram and twin, the better we can reconstruct the phase.

value resets. From the graph, we can notice that the best SNR occurs at around $R = 3.5$. While the optimal value depends on the noise and frequency content of the scene, we observed that it is always around $R = 4.0$, and the SNR difference between the optimal and $R = 4.0$ is not significant.

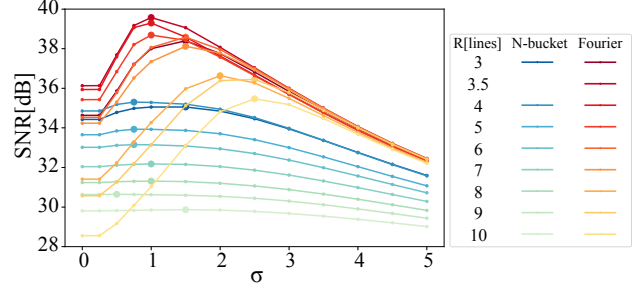


Figure 9. For the snapshot image, instead of the proposed Fourier technique, we can also use the standard N-bucket technique by using neighboring rows as a proxy for the remaining phases. However, the N-bucket technique performs significantly worse compared to the Fourier technique at all the phase variation rates and cannot handle non-integer values of R .

cant. At $R = 4$, the overlap between the ToF hologram and the twin is minimal, and $\sigma = 1$ corresponds to suppressing half of the higher frequency content; hence, the optimal SNR is always around $R = 4$. When $R = 0$, the twin and hologram are inseparable and the phase cannot be correctly recovered as shown in Fig. 8. In the rest of the paper, we use the phase variation rate of $k = \pi/2$ and $\sigma = 1$.

Comparison with N-bucket technique Instead of the Fourier reconstruction method detailed in Sec. 4, we can also use the standard N-bucket technique presented in Sec. 3 by grouping R -rows with phase variations $\theta = \{0, 2\pi/R, 4\pi/R, \dots, (R-1)/R\}$. Note that the N-bucket technique does not decrease the number of effective rows, as we can always group $R/2$ rows around the current row to get all the phase variations, in a sliding window manner. However, unlike the Fourier reconstruction method, the N-bucket technique works only for integer values of R .

In Fig. 9, we compare the N-bucket technique with the Fourier technique and observe that the Fourier technique uniformly works better than the standard N-bucket technique for snapshot phase reconstruction.

Comparison with conventional operation We compare the proposed snapshot technique with the conventional four-bucket method for the same exposure duration. We capture the snapshot image with an exposure duration of T where $T = 100 \mu s$ while the four quadrature measurements are captured with $T/4$. In Fig. 10, we compare phase reconstruction results for both techniques. For a fair comparison, we have prefiltered the phasors for both the conventional and snapshot techniques. Our method has an SNR of 29.86 dB for phase reconstruction, slightly higher than 28.56 dB for the conventional method but our technique performs uniformly better than the conventional technique for all σ

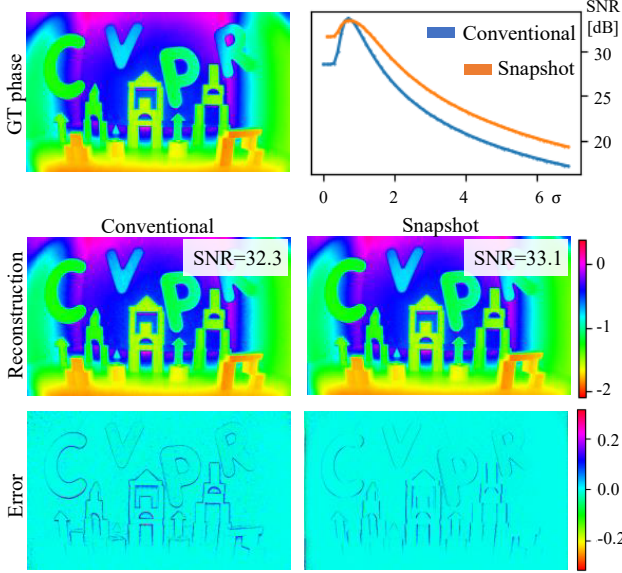


Figure 10. **Comparisons against quadrature reconstruction.** For the same total exposure duration and optimal kernel size, our method performs similarly or marginally better, compared to the conventional quadrature measurement technique. Importantly, however, our technique requires $4\times$ less bandwidth.

values. While the proposed technique performs slightly better in this case, the required bandwidth is $4\times$ smaller than the conventional method as our technique requires only one image.

Optimal phase variation direction Rotating the Fourier shift direction, implementable by rotating the camera along its optical axis, can also reduce the aliasing artifacts depending on the edges in the scene. We acquire the optimal result when the majority of the high-frequency edges are parallel to the direction of the phase shift (which is the same as being perpendicular to the phase variation direction). In Fig. 11, we demonstrate the result of reconstructing a scene made up of rectilinear block objects. To ensure the scene stays the same when rotated, we apply a circular mask when comparing the result with the ground truth and compute the SNR. As the graph shows, rotating the input at 75° yields the best result, since only a few rectilinear edges align with the vertical axis, the same orientation for the phase shift.

7. Conclusions

Inspired by off-axis holography, we proposed a snapshot Lidar using CW-ToF cameras, which captures amplitude and depth using a single image. We showed how defining a time-of-flight hologram and using rolling-shutter cameras allows us to translate off-axis principles to CW-ToF cameras. Extensive experiments with our lab prototype demon-

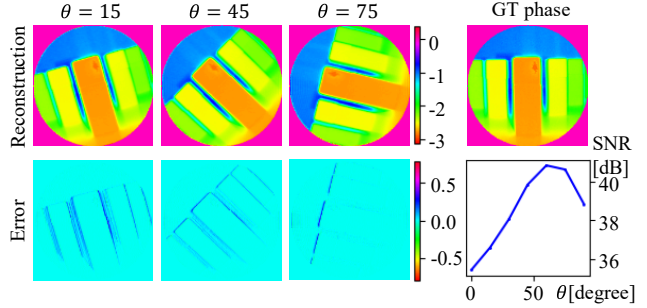


Figure 11. **Improving reconstruction by rotating the camera.** Rotating the rolling shutter or phase variation direction improves the reconstruction quality. When the phase variation direction is perpendicular to the dominant edge direction, then we attain the highest SNR. We change the phase variation direction by simply rotating the camera.

strated that our snapshot imaging approach performs as well as conventional quadrature measurement-based approaches, while requiring $4\times$ fewer measurements.

As mentioned in the related work, continuous-wave time-of-flight cameras saw a proliferation of applications in the last decade. Translating the snapshot techniques proposed in this paper for non-sinusoidal codes [10], Doppler time-of-flight imaging [13, 14], frequency-based light transport probing [25], and time-gating [28], while non-trivial, can significantly improve all the applications.

Our approach relies on spatial modulation to compute the ToF hologram, which required pre-filtering to avoid aliasing. As part of future work, we will explore machine learning-based techniques to relax the aliasing requirements. Data-free techniques such as deep image priors [29], and generative approaches such as diffusion models [5] will enable us to solve the inverse problem more robustly, and without explicit anti-aliasing constraints.

We have translated snapshot off-axis imaging techniques to CWToF imaging. However, a wide variety of imaging ideas including high dynamic range imaging, light-field imaging, polarization imaging, and spectral imaging today are implemented with assorted pixels, which are expensive, prone to aliasing, or require custom demosaicking algorithms. Extending off-axis principles to create assorted frequency-embedding is an interesting future direction.

8. Acknowledgements

The authors thank Jeremy Klotz for his help with the hardware prototype. The research is partially supported by a Burke Research Initiation Award.

9. Errata

Eq. (1) lacks the imaginary unit j in the expressions for the complex light waves and should be corrected as:

$$\begin{aligned} I &= |E_s(x, y)e^{-j\phi_s(x, y)} + E_r e^{-j k x \sin(\theta)}|^2 \\ &= |E_s(x, y)|^2 + |E_r|^2 + E_r^* E_s(x, y)e^{-j(k x \sin(\theta) + \phi_s(x, y))} \\ &\quad + E_r E_s^*(x, y)e^{j(k x \sin(\theta) + \phi_s(x, y))} \end{aligned}$$

with the changes in red. This error is present only in Eq. (1), and does not affect our derivation of the ToF hologram, as j was correctly included in relevant subsequent equations.

References

- [1] Supreeth Achar, Joseph R. Bartels, William L. 'Red' Whitaker, Kiriakos N. Kutulakos, and Srinivasa G. Narasimhan. Epipolar time-of-flight imaging. *ACM TOG*, 36(4), 2017. 2, 3
- [2] Manuel Ballester, Heming Wang, Jiren Li, Oliver Cossairt, and Florian Willomitzer. Single-shot ToF sensing with sub-mm precision using conventional CMOS sensors. *arXiv preprint arXiv:2212.00928*, 2022. 3
- [3] Manuel Ballester, Heming Wang, Jiren Li, Oliver Cossairt, and Florian Willomitzer. Single-shot tof sensing with sub-mm precision using conventional cmos sensors. *arXiv preprint arXiv:2212.00928*, 2022. 3
- [4] Clara Callenberg, Felix Heide, Gordon Wetzstein, and Matthias B Hullin. Snapshot difference imaging using correlation time-of-flight sensors. *ACM TOG*, 36(6):1–11, 2017. 3
- [5] Hyungjin Chung, Byeongsu Sim, Dohoon Ryu, and Jong Chul Ye. Improving diffusion models for inverse problems using manifold constraints. *NeurIPS*, 35:25683–25696, 2022. 8
- [6] Miguel Heredia Conde. A material-sensing time-of-flight camera. *IEEE Sensors Letters*, 4(7):1–4, 2020. 1
- [7] Etienne Cuche, Pierre Marquet, and Christian Depeursinge. Spatial filtering for zero-order and twin-image elimination in digital off-axis holography. *Applied optics*, 39(23):4070–4075, 2000. 3
- [8] Jan Fischer, Benjamin Huhle, and Andreas Schilling. Using time-of-flight range data for occlusion handling in augmented reality. *IPT/EGVE*, 109116, 2007. 1
- [9] Mohit Gupta, Shree K Nayar, Matthias B Hullin, and Jaime Martin. Phasor imaging: A generalization of correlation-based time-of-flight imaging. *ACM Transactions on Graphics (ToG)*, 34(5):1–18, 2015. 2, 3, 4
- [10] Mohit Gupta, Andreas Velten, Shree K Nayar, and Eric Breibach. What are optimal coding functions for time-of-flight imaging? *ACM TOG*, 37(2):1–18, 2018. 8
- [11] Felipe Gutierrez-Barragan, Syed Azer Reza, Andreas Velten, and Mohit Gupta. Practical coding function design for time-of-flight imaging. In *CVPR*, 2019. 3
- [12] Miles Hansard, Seungkyu Lee, Ouk Choi, and Radu Patrice Horaud. *Time-of-flight cameras: principles, methods and applications*. Springer Science & Business Media, 2012. 2
- [13] Felix Heide, Wolfgang Heidrich, Matthias Hullin, and Gordon Wetzstein. Doppler time-of-flight imaging. *ACM TOG*, 34(4):1–11, 2015. 1, 8
- [14] Yunpu Hu, Leo Miyashita, and Masatoshi Ishikawa. Differential frequency heterodyne time-of-flight imaging for instantaneous depth and velocity estimation. *ACM Transactions on Graphics (TOG)*, 42(1):1–13, 2022. 8
- [15] Suren Jayasuriya, Adithya Pediredla, Sriram Sivaramakrishnan, Alyosha Molnar, and Ashok Veeraraghavan. Depth fields: Extending light field techniques to time-of-flight imaging. In *Intl. Conf. 3D Vision*, 2015. 2
- [16] Achuta Kadambi, Refael Whyte, Ayush Bhandari, Lee Streeter, Christopher Barsi, Adrian Dorrington, and Ramesh Raskar. Coded time of flight cameras: sparse deconvolution to address multipath interference and recover time profiles. *ACM TOG*, 32(6):1–10, 2013. 3
- [17] Jongho Lee and Mohit Gupta. Stochastic exposure coding for handling multi-tof-camera interference. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 7880–7888, 2019. 4
- [18] Jongho Lee and Mohit Gupta. Mitigating ac and dc interference in multi-tof-camera environments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2023. 4
- [19] Emmett N Leith and Juris Upatnieks. Reconstructed wavefronts and communication theory. *JOSA*, 52(10):1123–1130, 1962. 3
- [20] Fengqiang Li, Huaijin Chen, Adithya Pediredla, Chiakai Yeh, Kuan He, Ashok Veeraraghavan, and Oliver Cossairt. CS-ToF: High-resolution compressive time-of-flight imaging. *Optics express*, 25(25):31096–31110, 2017. 2
- [21] Fengqiang Li, Florian Willomitzer, Prasanna Rangarajan, Mohit Gupta, Andreas Velten, and Oliver Cossairt. SH-ToF: Micro resolution time-of-flight imaging with superheterodyne interferometry. 2018. 3, 4
- [22] Fengqiang Li, Florian Willomitzer, Muralidhar Madabhushi Balaji, Prasanna Rangarajan, and Oliver Cossairt. Exploiting wavelength diversity for high resolution time-of-flight 3d imaging. 43(7):2193–2205, 2021. 3
- [23] Melexis. Melexis 75027 application note. <https://media.melexis.com/-/media/files/documents/product-flyers/mlx75027-product-flyer-melexis.pdf>. 4
- [24] Nikhil Naik, Achuta Kadambi, Christoph Rhemann, Shahram Izadi, Ramesh Raskar, and Sing Bing Kang. A light transport model for mitigating multipath interference in time-of-flight sensors. In *CVPR*, pages 73–81, 2015. 2, 3
- [25] Matthew O'Toole, Felix Heide, Lei Xiao, Matthias B Hullin, Wolfgang Heidrich, and Kiriakos N Kutulakos. Temporal frequency probing for 5D transient analysis of global light transport. *ACM Transactions on Graphics (ToG)*, 33(4):1–11, 2014. 2, 8
- [26] Matthew O'Toole, Felix Heide, Lei Xiao, Matthias B Hullin, Wolfgang Heidrich, and Kiriakos N Kutulakos. Temporal frequency probing for 5d transient analysis of global light transport. *ACM TOG*, 33(4):1–11, 2014. 4
- [27] Hoover Rueda-Chacon, Juan F Florez-Ospina, Daniel L Lau, and Gonzalo R Arce. Snapshot compressive tof+ spectral

imaging via optimized color-coded apertures. 42(10):2346–2360, 2019. 2

- [28] Ryuichi Tadano, Adithya Kumar Pediredla, and Ashok Veer-araghavan. Depth selective camera: A direct, on-chip, programmable technique for depth selectivity in photography. In *ICCV*, 2015. 4, 8
- [29] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Deep image prior. In *CVPR*, 2018. 8
- [30] Mian Wei, Navid Sarhangnejad, Zhengfan Xia, Nikita Gu-sev, Nikola Katic, Roman Genov, and Kiriakos N Kutulakos. Coded two-bucket cameras for computer vision. In *ECCV*, 2018. 3
- [31] Di Wu, Andreas Velten, Matthew O’Toole, Belen Masia, Amit Agrawal, Qionghai Dai, and Ramesh Raskar. Decomposing global light transport using time of flight imaging. *IJCV*, 107:123–138, 2014. 2