SH-ToF:
Micro Resolution Time-of-Flight Imaging with Superheterodyne Interferometry

Fengqiang Li[1] Florian Willomitzer[1] Prasanna Rangarajan[2] Mohit Gupta[3]
Andreas Velten[3] Oliver Cossairt[1]
[1]Northwestern University [2]Southern Methodist University [3]University of Wisconsin-Madison
fengqiang.li@u.northwestern.edu

Abstract

Three dimensional imaging techniques have been widely used in both industry and academia. Time-of-flight (ToF) sensors offer a promising method of 3D imaging due to compact size and low complexity. However, state-of-the-art ToF sensors only have depth resolutions of centimeters due to limitations in the modulation frequencies that can be used. In this paper, we propose a technique to generate modulation frequencies as high as 1 THz using optical superheterodyne interferometry. Our proposed system provides great flexibility in imaging range and resolution. We experimentally demonstrate an increase in depth resolution by an order of magnitude relative to currently available commercial ToF cameras.

1. Introduction

Continuous-wave ToF sensors are fast emerging as low-cost distance/3D shape measuring devices [26, 18, 28]. The basic operating principle of these sensors is to illuminate the scene with a light source whose intensity is modulated over time, for example, as a sinusoid [19]. The detector pixels in these devices behave as homodyne receivers [19, 8] that accumulates a charge proportional to the phase-shift between the emitted light, and the radiance received at the sensor. For each pixel, the scene distance z can be estimated from the measured phase-shift ϕ_t as:

$$z = \frac{1}{2}\frac{c}{f_t}\frac{\phi_t}{2\pi} \tag{1}$$

where f_t is the temporal modulation frequency of the light source, and c is the speed of light. The range resolution δz is inversely proportional to the modulation frequency [19]:

$$\delta z \sim \frac{1}{f_t} \tag{2}$$

The use of silicon based detectors limits the temporal frequencies to tens or hundreds of MHz. This in turn limits the achievable range resolution of current ToF cameras to centimeters [18, 28].

Can this technical restriction be bypassed? This limitation can be overcome by borrowing ideas from interferometry [1, 23, 21], which provides considerably higher range resolution (order of microns) by carefully engineering the spatio-temporal correlation properties of the light source. The correlation is exploited in identifying the exact distance to target by comparing the travel time of the reflected light from the object with a reference beam. The difference in travel times manifests as a phase shift in the spatial pattern observed at the detector.

The notion of comparing the travel time of the return signal to a reference signal is analogous to the concept of operation of ToF cameras. The difference lies in the fact that the reference signal in ToF cameras is an electronically generated RF signal that is limited to tens or hundreds of MHz. In contrast, the reference signal used in interferometry is an electromagnetic field that oscillates at optical frequencies in excess of 100 THz. The use of such high frequencies provides exceptionally high range resolution. However, the unambiguous measurement range is restricted to the optical wavelength (order of microns). This is because the measured phase wraps around after 2π, and scene points that are exactly half wavelength away have the same measured phase. This problem, called phase wrapping, is more severe for higher modulation frequencies; higher the modulation frequency, lower the unambiguous depth range.

A second issue arising from the use of coherent light in interferometry is the mottled/grainy appearance of the image of the object (called speckle). Speckle arises from the stochastic character of the roughness of real world object and surfaces at scales comparable to the wavelength [11].

The present work seeks to bridge the divide in the range resolution of current ToF cameras and interferometers (see Fig. 1(d)). As a first step towards achieving this goal, we introduce the design and development of a superheterodyne interferometry (SH-ToF) ranging device with submillimeter resolution and an unambiguous measurement range of centimeters. The imager concept is illustrated in Figure 1(c).

The competing requirements of increased range reso-

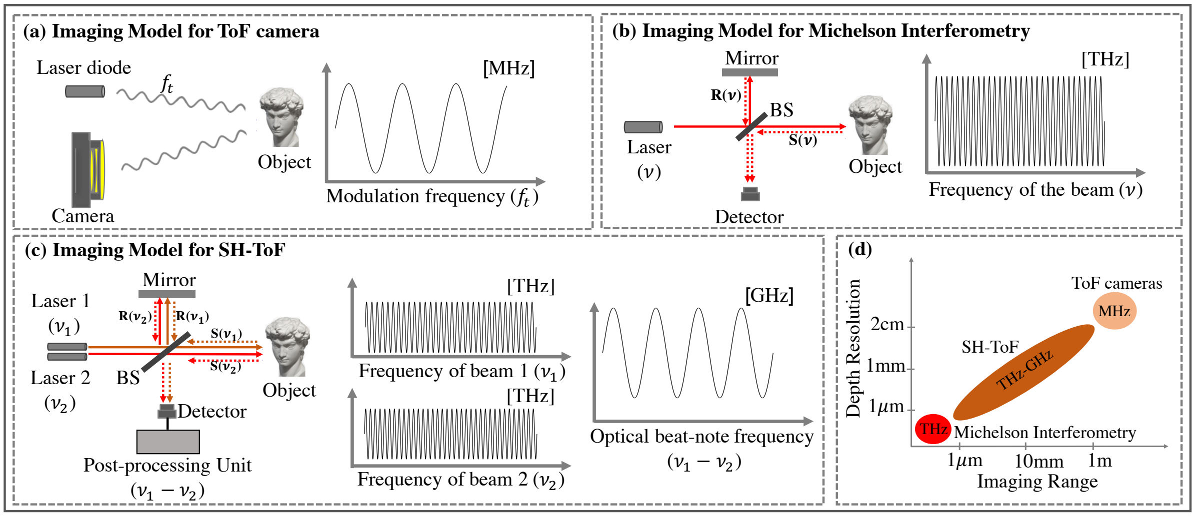

Figure 1: (a). **Imaging model for ToF cameras**: a RF (f_t) signal modulates the amplitude of the illumination beam from the laser diode and the detector with an arbitrary phase shift. (b). **Imaging model for Michelson Inteferometry**: the beam ($Ee^{i2\pi\nu t}$) from the laser is the illumination in the system, and the modulation frequency is the optical frequency ν. (c). **Imaging model for our proposed setup**: two lasers with frequencies of ν_1 and ν_2 are used as illumination sources. A 'new' optical frequency of $(\nu_1 - \nu_2)$ is generated after the post-processing unit, which enables micro depth resolution with macro imaging range. (d). Depth resolutions and image ranges for these three different imaging models. BS: beam splitter. R(ν): beam reflected from the mirror. S(ν): beam reflected from the object.

lution and large unambiguous range are simultaneously accommodated by using two coherent light sources with closely spaced optical frequencies (say ν_1 and ν_2 THz). The notion is borrowed from the concept of operation of a superheterodyne interferometer [5]. The redundancy in the interference patterns associated with the two wavelengths may be exploited to electronically (or computationally) generate an interference pattern at the synthetic wavelength ($\Lambda = c/|\nu_1 - \nu_2| = \lambda_1\lambda_2/|\lambda_1 - \lambda_2|$), which is considerably larger than the individual wavelengths.

A number of superheterodyne systems have been proposed and implemented using two separate lasers with fixed optical frequencies. Ranging has been demonstrated with unambiguous depth ranges of tens of mm and sub-mm depth resolution [5, 25, 7]. In this paper, we propose a new superheterodyne system implemented using tunable lasers, providing the following advantages in 3D imaging:

- **Synthetic wavelength phase measurement.** The phase measured by the device is determined by the synthetic wavelength and not the optical wavelengths of the two sources. However, the lateral resolution of the proposed imager is still limited by the optical wavelengths (much smaller).
- **Flexible tradeoff between range and resolution.** The flexibility in the selection of the depth resolution (from 10's of microns to millimeters) and unambiguous range (from millimeters to meters) afforded by tuning the difference in the optical frequencies of the two sources over the GHz to THz interval.
- **Insensitivity to environmental fluctuations.** While interferometers are typically sensitive to small fluctuations caused by air currents, the proposed superheterodyne interferometer is only sensitive to fluctuations exceeding in the order of the synthetic wavelengths, which is of the order of millimeters.

The use of superheterodyne interferometry principles for high resolution ranging in a reduced form factor has recently been demonstrated by Li [22]. The proposed work expands the scope of previous work [22], by making the following contributions:

- **Tunability of depth resolution and depth range**. We experimentally demonstrate the flexibility in the selection of depth resolution (from tens of microns to millimeters) and unambiguous range (from millimeters to meters) by tuning the difference in the optical frequencies of the two laser sources.
- **Full field 3D scanning of optically rough surfaces**. To the best of our knowledge, the experiments disclosed in this contribution represent the first documented demonstration of full-field 3D scanning using superheterodyne interferometry.
- **Error analysis**. We analyze factors affecting system performance, including laser frequency drift, speckle noise and photon noise.

2. Related Work

Most work in ToF imaging deals with overcoming the limited spatial resolution of ToF sensors [16, 29, 20, 30], and mitigating errors due to multipath interference [10, 17, 15, 24, 27]. Despite advances, the range resolution of current ToF sensors is fundamentally limited by the poor temporal response of silicon based ToF pixels at modulation frequencies in excess of 100 MHz [19, 13].

In a recent paper [22], a ToF imaging system based on SHI has been introduced working at GHz modulation frequencies. However, its performance is limited by the vibration of the non-polarization fiber and the non-linearity introduced by the applied electro-optical modulator (EOM). Kadambi [14] recently proposed a framework for high-resolution depth estimation using a stack of Mach-Zehnder interferometers and GHz modulation, while also demonstrating immunity to vibrations at optical scales. These works, however, provide simulation results and limited experimental evidence using objects and surfaces that are relatively smooth at optical scales. By combining spot illumination of the object and SNR filtering, we are for the first time able to use SHI to recover the shape and geometry of optically rough objects such as a plaster bust of David.

The work most similar to ours is by Fercher [6], who use a dual detector heterodyne interferometer to simultaneously measure a ToF signal at two different optical frequencies. The two measurements are then combined to compute the phase relative to a synthetic wavelength. Fercher also explores the effect of 3D scanning rough surfaces with a dual wavelength interferometer. However, our work differs in that we use a super-heterodyne principle to sense two wavelengths using a single detector. This results in two main differences: 1) Our technique is amenable to detection of two wavelengths with slight shifts in frequency (e.g., 1GHz to 1THz), and 2) Because our technique uses only a single detector for both wavelengths, snapshot (non-scanning) acquisition is possible using a focal plane array.

3. SH-ToF imaging system

A schematic of the proposed SH-ToF imager is shown in Fig. 2. Two lasers with slightly different wavelengths λ_1, λ_2 are utilized. As is common in interferometry, the light from each laser is fed into a beam splitter, to produce the sample beam (marked with red lines) and the reference beam (green lines). The sample beams from the two lasers are combined using a beam splitter (B_1) and directed towards the object. Likewise, the reference beams from the two lasers are combined using another beam splitter (B_2) and directed towards the photo-detector. The reference arms are additionally equipped with an acousto-optic modulator (AOM) that up-shifts the optical frequency by a prescribed amount (f_{m1} for ν_1, and f_{m2} for ν_2). The benefit of the up-shift will be made apparent in a subsequent paragraph.

The focusing optic Lens in Fig. 2 serves the dual purpose of illuminating a single spot on the object and imaging the object onto the detector. The two-axis galvo mirror system in Fig. 2 aids in scanning the focused sample beams across the object surface, and re-directing the backscattered light from the object towards the collection optic Lens.

The backscattered light from the object inteferes with the reference beams, at the respective wavelengths, due in large part to the increased coherence length of our lasers. An avalanche photo diode (APD) records the instantaneous irradiance of the detected inteference pattern. The APD readout is digitized with a data acquisition (DAQ) card. A post-processing unit recovers the measurements at the 'synthetic' frequency ($\nu_1 - \nu_2$).

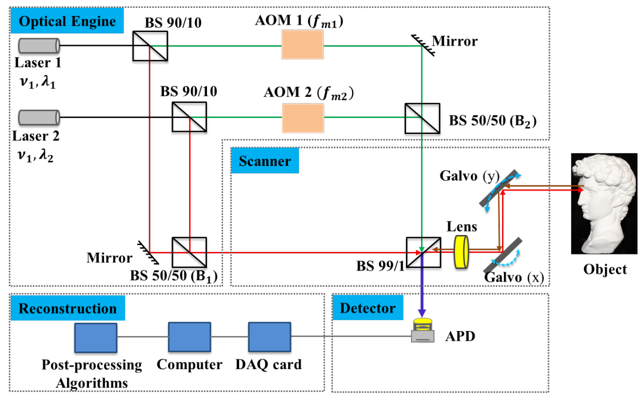

Figure 2: **Schematic of the proposed SH-ToF system.** Two lasers are used simultaneously as the light source in the system. Red lines: sample beams. Green lines: reference beams. Brown lines: beam reflected from the object. Purple lines: interference beams. 90/10: 90% of the beam directs to sample arm and the left 10 % beam goes to reference arm. 99/1: 99% of the beam reflecting from the object (brown lines) is directed to APD, and 1 % of the reference beam (green lines) goes to APD.

3.1. Theory of SH-ToF

The mathematical principles underlying the operation of the proposed Superheyerodyne ToF imager are described below. In the interest of simplicity, we impose the following restrictions

- the two laser sources emit linearly polarized narrowband light [1] with center wavelengths λ_1 and λ_2 respectively.
- light transport through the scene is adequately described by scalar field propagation.
- the blurring associated with the illumination and imaging optics may be disregarded.
- the effect of environmental vibration between sensor and object may be adequately modeled using a velocity-based vibration model [9].

Bearing in mind these restrictions, we can derive an expression for the instantaneous irradiance at the APD, which is disclosed below[1]:

$$I(t) = a_0 + a_1 \cos\left(\frac{Vt}{\lambda_1} + \frac{4\pi L}{\lambda_1} - 2\pi f_{m1} t\right) + a_2 \cos\left(\frac{Vt}{\lambda_2} + \frac{4\pi L}{\lambda_2} - 2\pi f_{m2} t\right) \quad (3)$$

where a_0, a_1, and a_2 are scalar constants. V represents the vibration velocity. L is the physical optical path difference (OPD) between the sample and reference arms. Please note that the optical path in the sample arm encapsulates the combined effect of macroscopic depth variations associated with topographic changes in the object and microscopic height variations due to surface roughness of the object. The phase variations associated with the macroscopic variations are of interest from the standpoint of ranging. The phase variations induced by the surface roughness are a source of measurement noise (speckle)

Inspection of Eq. 3 confirms that the APD irradiance is a superposition of sinusoids at the two AOM frequencies f_{m1}, f_{m2} respectively. The time-independent phase shift $\frac{4\pi L}{\lambda_{1,2}}$ associated with each sinusoid encodes the distance to the object, albeit with a 2π phase ambiguity. The ambiguity is resolved by computationally interfering the sinusoids at the frequency f_{m1} and f_{m2}. To this end, the APD output is digitized using a high-speed DAQ whose sampling rate exceeds $2 \times \max(f_{m1}, f_{m2})$. A squaring operation applied to the digitized signal yields a time varying signal that is comprised of multiple frequency components at f_{m1}, f_{m2}, $(f_{m1} - f_{m2})$, $(f_{m1} + f_{m2})$, $2f_{m1}$, and $2f_{m2}$ Hz.
A band-pass filter is then used to pick off the beat frequency component associated with the frequency difference $(f_{m1} - f_{m2})$. The expression for the filtered signal is disclosed below:

$$B(t) = m_1 \cdot \cos\left[\frac{Vt}{\Lambda} + \underbrace{\frac{4\pi L}{\Lambda}}_{\Phi(\mathrm{L})} - 2\pi(f_{m1} - f_{m2})t\right] + m_2 \quad (4)$$

where $m_1 = a_1 a_2$, and m_2 is a constant, and $\Lambda = c/|\nu_1 - \nu_2|$ is the synthetic wavelength.

It is evident that the filtered signal behaves as a sinusoid at the beat frequency $\Delta f = (f_{m1} - f_{m2})$. The first term inside the cosine function in Eq. 4 represents a random phase delay due to environmental vibrations. Because the phase delay is measured relative to the synthetic wavelength [2] and not the optical wavelength, the term will be largely negligible. This results in the expression for the filtered signal shown below:

$$B(t) \approx m_1 \cdot \cos\left[\frac{4\pi L}{c} \underbrace{(\nu_1 - \nu_2)}_{\text{Optical beat-note frequency}} - 2\pi \underbrace{(f_{m1} - f_{m2})}_{\text{AOM beat-note frequency}} t\right] + m_2 \quad (5)$$

A comparison of the aforementioned expression to the instantaneous irradiance recorded by a detector pixel in a ToF camera [19], suggests that the **proposed imager behaves as a ToF camera with a modulation frequency that matches the optical beat-note (synthetic) frequency** $(\nu_1 - \nu_2)$**, of our interferometer**. Tuning the laser wavelengths allows us to realize synthetic frequencies in the GHz to THz range.

The depth of the object may be recovered from the time independent phase shift associated with the beat frequency Δf, as shown below:

$$\Phi(L) = \mathrm{atan2}\{\mathrm{imag}\left(\mathscr{F}\{B(t)\}_{\Delta f}\right), \mathrm{real}\left(\mathscr{F}\{B(t)\}_{\Delta f}\right)\} \quad (6)$$

By careful selection of the laser wavelengths, the phase ambiguity in the depth measurement may be avoided, thereby increasing the unambiguous range of our ToF imager.

Additionally, the albedo of the object may be recovered from the filtered signal $B(t)$. It is observed that the amplitude m_1 of the sinusoidal component at the beat frequency Δf Hz is related to the object albedo, as follows:

$$m_1 = a_1 a_2 = \kappa \cdot \beta^2 \quad (7)$$

where the term β represents the object albedo and κ represents for a scalar constant.

3.2. Imaging range and depth resolution of SH-ToF

Imaging range: As mentioned previously, the unambiguous measurement range of our ToF imager is restricted to the synthetic wavelength Λ, which is between 3-48 mm for the proof-of-principle imager described in this contribution. The limited unambiguous range will likely introduce phase wrapping artifacts when trying to capture topographic variations in larger objects. The problem may be mitigated by employing state-of-the-art phase unwrapping techniques from interferometry [2, 3], such as measurements at multiple (synthetic) frequencies. The latter approach is particularly well suited for operation with tunable laser sources.

Depth resolution: Inspection of Eq. 2 suggests that increased depth resolution may be achieved by increasing the modulation frequency. In an effort to analyze the theoretical upper-bound depth resolution in our SH-ToF, a numerical simulation is performed. In the simulation, different optical beat-note frequencies of 0.1 THz (corresponding Λ of 3mm), 25 GHz (12mm), 12.5 GHz (24mm), and 6.25 GHz (48mm) are used, which are generated from two swept-source lasers. We follow the theory of SH-ToF in section 3.1 and build a simulator. A 1 mm/s velocity-noise is added in the simulator. Depth resolutions with different signal to noise ratios (SNR) are simulated as shown in Fig. 3, where root mean squared errors (RMSE) between estimated phases and ground truth phase are quantified and compared, and RMSEs between estimated depths and ground truth depth are also compared. This simulation result provides the theoretical upper-bound depth resolution.

[1] For a detailed derivation of the mathematical model, please refer to the Supplementary Materials

[2] In our experiments, we chose Λ between 3mm and 48mm, see Section 4

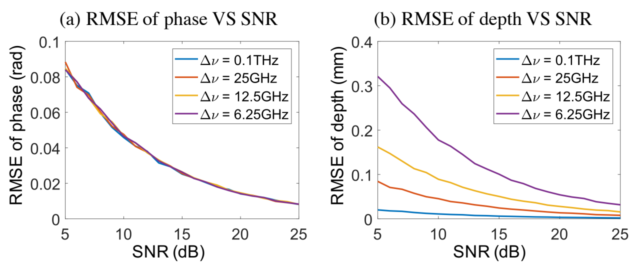

Figure 3: **Upper-bound depth resolutions with SH-ToF (Simulation Results):** (a). The RMSEs between estimated phases and the ground truth phase with different SNRs for different synthetic or optical beat-note frequencies. (b). The RMSEs between estimated depths and ground truth depth for different SNRs.

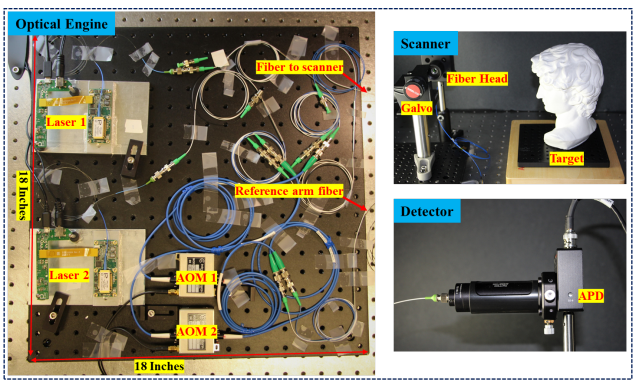

Figure 4: **Fiber-based experimental prototype:** All connections are fiber based except the scanner. Optical engine including lasers and AOMs is fit in a 18×18 inches aluminum board. Scanner contains a lens focusing the beam to the object and a Galvo. APD is used as the detector.

3.3. Proof-of-principle prototype

Figure 4 provides an illustration of our proof-of-principle ToF imager. All light paths in the system are routed through single-mode polarization maintaining fibers, with the exception of the light path to and from the object. A pair of inexpensive tunable laser sources (PPCL200, Pure Photonics) operating at $\lambda_1 \approx \lambda_2 \approx 1550\text{nm}$ drives the two arms of the Superheterodyne interferometer. The coherence length of each laser is about 3 kilometers which corresponds to a linewidth of 10 KHz. Light from each laser is split using a 90:10 fiber coupler (PC1550-90-APC, Thorlabs), where 10% of the laser output drives the reference arm and the remaining 90% drives the sample arm. The light in each of the reference arms is independently up-shifted at $f_{m1} = 40\text{MHz}$, $f_{m2} = 40.2\text{MHz}$ using two AOM's (FCM-401E, IntraAction Corp). The light from the sample/reference arms at the two wavelengths are combined with a 50:50 fiber coupler (PN1550R5A2, Thorlabs). The combined sample beam is directed towards the object through an assembly including a fiber circulator (CIR1550PM-APC, Thorlabs), a two-axis galvo mirror (GVS012, Thorlabs) and a focusing optic (AC080-020-C-ML, Thorlabs). The assembly serves the dual purpose of illuminating the object and re-directing the backscattered object light towards the detector. The interference of the backscattered object field and the reference is fiber-coupled to an APD (APD430C, Thorlabs) with maximum gain of 1.8×10^5 V/W and bandwidth of 400 MHz. The APD readout is digitized by a high-speed DAQ card (ATS9373, AlazarTech) with an analog input of 12-bit resolution.

4. Experiments and Results

This section presents a compilation of results from a series of ranging experiments conducted using our prototype ToF imager. The experiments are organized into three categories of increasing scene complexity: single-point measurement, line-profile measurement and full-field measurement. In each case, system performance is assessed for multiple synthetic frequencies ranging from 6.25 GHz to 100 GHz (see Tab. 1). For the case of full field measurement, a comparison with a commercially available ToF camera is included for reference.

Table 1: Laser-wavelengths-combinations (λ_1, λ_2) used in the shown experiments, leading to different sythetic wavelengths Λ, resp. synthetic frequencies $\Delta\nu$.

Set	1	2	3	4	5
λ_1 [nm]	1550	1550	1550	1550	1550
λ_2 [nm]	1550.8	1550.4	1550.2	1550.1	1550.05
Λ [mm]	3	6	12	24	48
$\Delta\nu$ [GHz]	100	50	25	12.5	6.25

4.1. Fixed-point Measurement

In an effort to quantify the upper-bound on the precision of our range measurements, we attempt to repeatedly (10,000 times) measure the depth of a fixed scene point at a standoff distance of 300 mm. The 2-axis galvo mirrors are held fixed in position during the course of the experiment. The beam from the fiber head is focused on a planar surface (cardboard) that is rough at optical wavelength scales, and consequently induces speckle artifacts. The measurement process is repeated for multiple synthetic frequencies (set 1, 3, 4, and 5 of Tab. 1) to help assess the dependence on synthetic frequency. The DAQ operates at 500M samples/sec for 0.2 seconds, collecting a total of 100M samples.

Results: Figure 5 plots the depth measurements for the various synthetic frequencies. For each acquisition, the corresponding depth values z are plotted.

The measurement precision in phase (given by $\delta\Phi$) and in depth (given by δz) are tabulated in Tab. 2. It is observed that $\delta\Phi$ is largely constant across synthetic wavelengths Λ, thereby suggesting that the depth precision δz

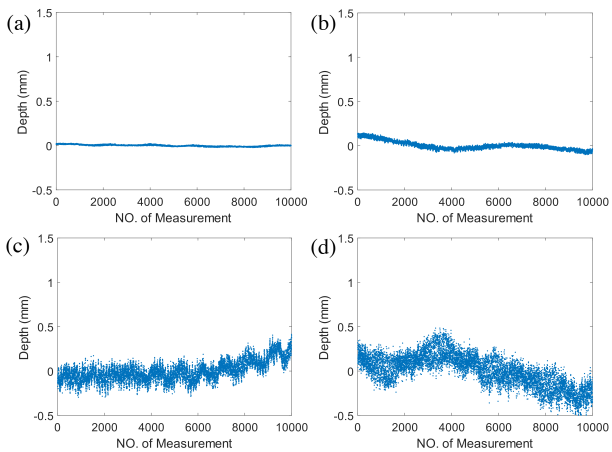

Figure 5: **SH-ToF for fixed-point measurement:** The depth estimations with 10,000 repeated measurements for optical beat-note frequency of 0.1 THz (a), 25GHz (b), 12.5 GHz (c), and 6.25 GHz (d). Each set of 10,000 repeated measurements is finished in 0.2 seconds.

improves with decreasing synthetic wavelength (or increasing synthetic frequency).

For these experiments, we measured a signal to background ratio of SNR $\approx$ 11 dB in our prototype imager. For this SNR, the experimentally observed precision is in agreement with the theoretical predictions disclosed in section 3.2 and Fig. 3(b).

The experimentally validated upper-bound on the range precision our prototype imager is consistent with theoretical predictions, in spite of the relatively long measurement durations (e.g., 0.2 seconds).

Table 2: Acquired precisions for phase - ($\delta\Phi$) and depth-precision (δz) using different synthetic wavelengths.

Set	**1**	**3**	**4**	**5**
$\Delta\nu$ [GHz]	100	25	12.5	6.25
Λ [mm]	3	12	24	48
$\delta\Phi$ [rad]	0.041	0.049	0.059	0.047
δz [mm]	0.009	0.047	0.114	0.179

Practical challenges: Most real world objects and surfaces exhibit height fluctuations at the microscopic scale and are consequently rough at the scale of the optical wavelengths considered in our experiments. This roughness combined with the increased coherence length of our sources introduces speckle artifacts (grainy appearance due to repeated constructive and destructive interference) in our range measurements. The speckle manifests as high contrast variations in the received signal strength for a surface with constant albedo and constant depth. As a result, the phase and depth measurement in the regions of destructive interference are unreliable. This suggests that the SNR of our ToF imager is fundamentally limited by speckle noise, and not shot noise as in traditional ToF cameras.

The impact of speckle on range measurements has been examined by Fercher [6], and they recommend making measurements in the vicinity of a speckle intensity maximum (regions of high SNR). It is made possible by sliding the object relative to the imager as we utilize for the fixed-point measurements in this section. However, this approach is infeasible in practice such as in following scanning measurements (sections 4.2 and 4.3). Therefore, we propose an alternative method with more details in the Supplementary Materials.

4.2. Line-scanning Measurement

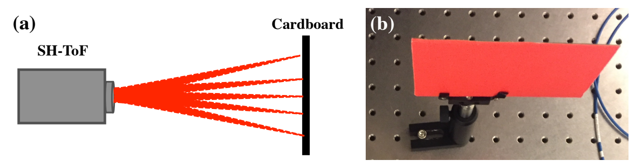

Figure 6: **SH-ToF for line scanning on a planar surface:** (a). The beam is focused on the object and scanned along a line on the planar surface with a Galvo scanner. (b). A cardboard is used as the planar surface.

The second set of experiments described in this section are quantitative characterizations of the precision in scanning measurements. Herein, we attempt to measure the depth of a fixed line on a planar surface positioned at a standoff distance of 300 mm (shown in Fig. 6). For the purpose of line-scan measurement, one of the mirrors in the 2-axis galvo mirrors is held fixed, while the other mirror is swept (by applying a sinusoidal drive at 10 Hz) over an angular range of ± 0.244 rad. The angular sweep roughly corresponds to a scan length of 8mm at standoff, and yields 2500 measurements over a 50 ms acquisition interval. The measurements are repeated across the multiple synthetic frequencies shown in Tab. 1(Set 1-5). It is expected that the measured phase will exhibit a 2π phase wrapping ambiguity for synthetic wavelengths smaller than the length of a scan-line at standoff (8mm in the present case).

For reasons stated previously, the precision of our line-scanning measurement is fundamentally limited by speckle.

Results: The results of the line-scan experiment are tabulated in Fig. 7 and Tab. 3. As expected, the measurement precision improves with increasing synthetic frequency.

The plots in second column of Fig. 7 illustrate the raw depth values as acquired by the SH-ToF imager. The phase wrapping at smaller synthetic wavelengths is evident in rows 1-3. The unwrapped depth maps are illustrated in column 3 of Fig. 7.

Inspection of the unwrapped depth profile in Fig. 7 confirms that the depth profile is roughly quadratic (see more in Supplementary), which is consistent with the increase in euclidean distance as the galvo mirror scans across the surface of the object.

We have experimentally validated the micro-resolution in the range measurements provided by our prototype imager, and additionally demonstrated the flexibility in the selection of depth resolution and unambiguous range by tuning the synthetic frequency.

Table 3: Precision of the line-scanning measurements

Set	**1**	**2**	**3**	**4**	**5**
$\Delta\nu$ [GHz]	100	50	25	12.5	6.25
δz [mm]	0.070	0.093	0.221	0.274	0.437

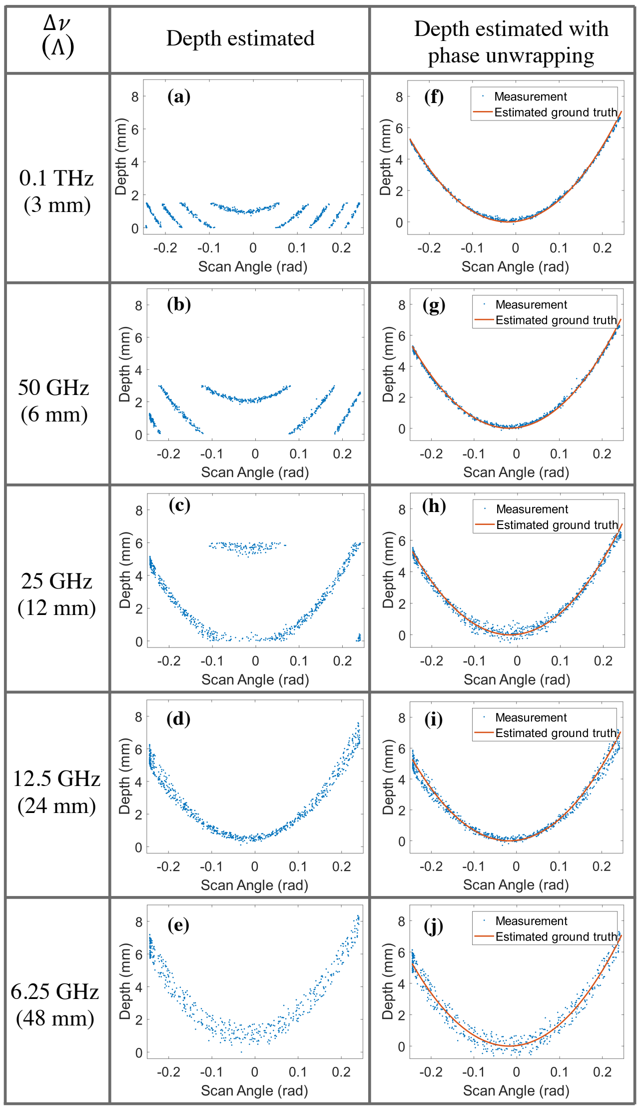

Figure 7: **Line-scanning measurements with different optical beat-note frequencies:** Original depth estimated (left column) and depth with phase unwrapping (right column) are shown for different optical beat-note frequencies. A ground truth (red lines) is estimated with described method. **Note:** Image ranges are half of the corresponding synthetic wavelengths due to the round trip in sample arm.

4.3. Full Field 3D Scanning

The experiments discussed thus far have focused on quantifying the precision of our depth measurements. We now proceed to demonstrate 3D scanning of complex objects such as the bust of David. For reference, we compare the performance of our prototype SH-ToF imager with that of a commercially available ToF camera (Texas Instrument OPT8241, 240×320 pixels).

A full field 3D scan is performed by raster scanning the laser spot in the horizontal (x) and vertical (y) directions, using the 2-axis galvo (see Fig. 8). For each y-position, a line-scan in the x-direction is used to obtain 2500 measurements over a 50ms interval. The process is repeated for 100 distinct y-positions, producing a point-cloud with 0.25 million independently measured points. The scan area is approximately $\Delta X \times \Delta Y = 80\text{mm} \times 40\text{mm}$ at the prescribed stand-off (300 mm). In the interest of brevity, we only report the measurements at a synthetic wavelength of $\Lambda = 48\text{mm}$ (Set 5).

It is evident from the line-scan depth profiles of Fig. 7 that the depth values reported by our ToF imager for a strictly planar surface, exhibit geometric distortion. The behavior is reminiscent of radial distortion in scanning systems. The distortion can be compensated by learning the 2D mapping between the measured and expected depth values of a planar surface, which is used to compensate for the depth distortion encountered when scanning arbitrary 3D objects. More details are in Supplementary Materials.

4.3.1 Folded Cardboard

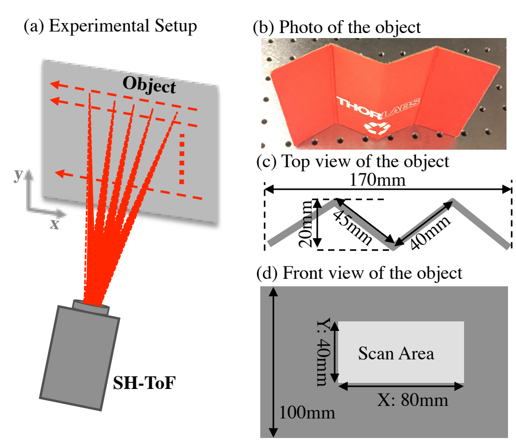

Figure 8: **SH-ToF for scanning a folded cardboard:** (a). Experimental setup. 3D scanning procedure: the beam is scanning along one red dash line in x axis, and then move in y axis to another red dash line and scan. (b). Photo of a folded cardboard. (c). The schematic of the object from top view. (d). The schematic of the object from the front view.

The first 3D scanning experiment described herein demonstrates the ability to recover the geometry of a planar object with multiple folds, such as the one shown in Fig. 8.

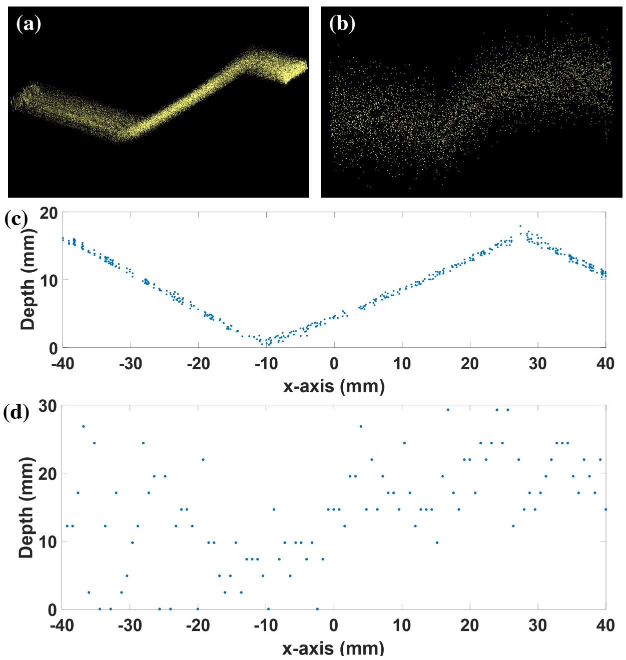

Figure 9: (a). Point cloud results with the proposed SH-ToF (roughly top view). (b). Point cloud results with a state-of-the-art ToF sensor (roughly top view). (c). A cross-line profile of the point cloud in (a). (d). A cross-line profile along the point cloud in (b).

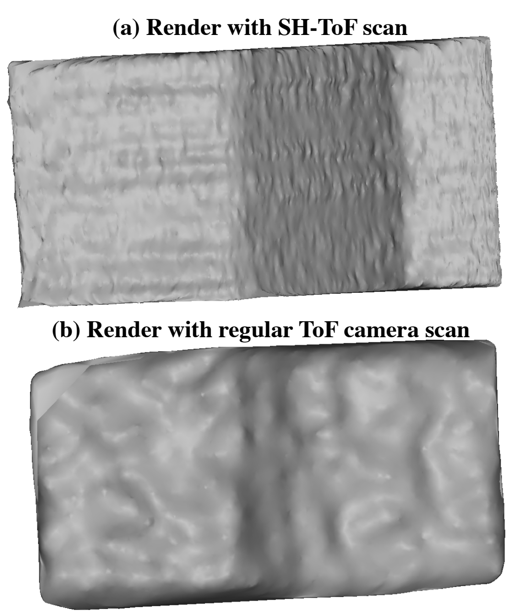

Figure 10: **Rendering results:** (a). Rendering result with SH-ToF point clouds (Front view). (b). Rendering result with ToF camera point clouds (Front view).

Results: The raw point cloud acquired with our prototype as well as the reference ToF camera are shown in Fig. 9 (a) and (b). A slice through each point-cloud (Fig. 9 (c) and (d)) demonstrates the measurement fidelity of the proposed ToF imager. It is evident the proposed ToF imager is a significant improvement over the state of the art, in that it can recover high quality range maps while resolving slope discontinuities. The acquired point clouds are rendered with a free software *'CloudCompare'* [4]. Identical parameters are used for the rendering of each pointcloud. Rendering of the result from a novel viewpoint is shown in Fig. 10.

4.3.2 Plaster Bust

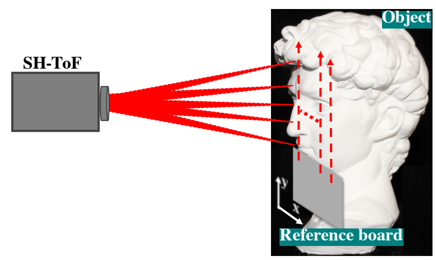

Figure 11: **Scanning setup for the plaster bust:** A plaster bust of David is used as the target, and a planar board under the bust serves as the reference board. The beam scans in y axis, and moves along x axis for repeating. Each scan starts from the planar board.

The final experiment described in this contribution is a 3D scan of a plaster bust of David. A schematic of the experimental setup is shown in Fig. 11. The estimated phases between two line scans may shift due to the laser frequency drift (see more detials in Supplementary). In order to align different line scans, a planar surface positioned below the bust of David serves as a reference surface for registration of each line scan. Alignment algorithms such as the interative closest points (ICP) algorithm [12] can potentially provide superior registration and relax the requirement for a reference surface.

Results: Figure 12 includes illustrations of the bust of David, and rendered views of the 3D scan acquired with our prototype and the reference ToF camera. It is evident that the proposed SH-ToF imager (Fig. 12(b1,2)) provides 3D depths maps with fine spatial detail. Closer inspection of the face of David (the eye area, Fig. 12 close-up images) reveals the vast difference in the resolution of the proposed imager and a state-of-the-art ToF camera.

5. Discussion

The discussion thus far has demonstrated the feasibility of using tunable sources and the principle of Superheterodyne Interferometry to realize a ToF imager with unprece-

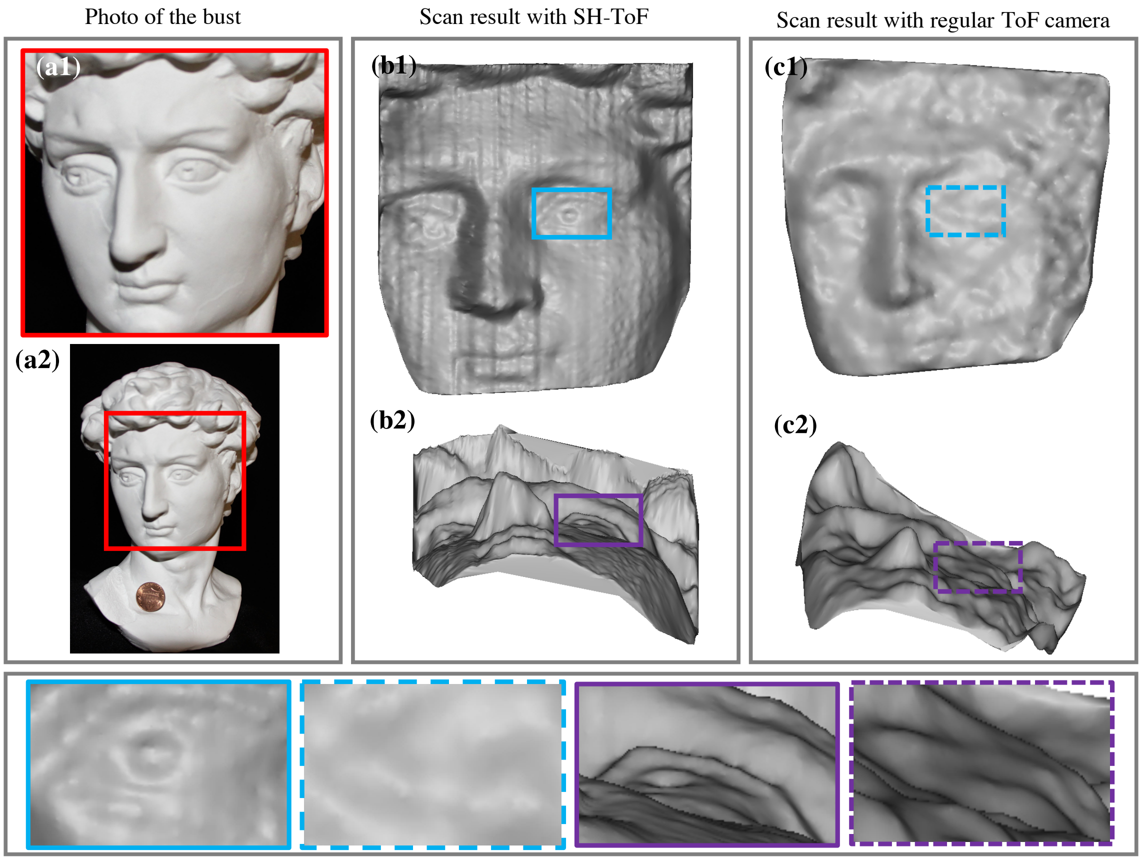

Figure 12: **Scanning result for the plaster bust:** (a2) shows the photo of the bust with scanning area of (a1). (b1 and b2). Scanning result with the proposed SH-ToF in the front view and side view respectively. (c1 and c2). Corresponding scanning results with a regular ToF camera in the front view and the top view respectively. Close-ups images in the bottom show more details of the scanning results with these two cameras. A US one-cent coin is used as a reference shown in (a2). Our SH-ToF 3D scanning result clearly demonstrates far superior depth resolution.

dented depth resolution. Replicating our success requires a comprehensive understanding of the issues that affect the fidelity of SH-ToF measurements. The present section is devoted to a discussion of these issues.

5.1. Noise

The dominant source of noise in traditional ToF cameras is ‘photon noise’ arising from the statistical variation in the arrival of photons at each detector pixel. SH-ToF imagers are no different in that it they are also susceptible to ‘photon noise’. However, the interferometric nature of our ToF imager means that we incur additional noise penalties due to speckle, laser wavelength (frequency) drift and environmental vibrations. These additional sources of noise are the price we pay for a substantial improvement in range resolution. Please see the Supplementary Material for a detailed analysis of these sources of noise.

5.2. Future Prospects

Our current approach to 3D scanning closely resembles LIDAR in that a dense point-cloud representation of the scene is assembled by raster scanning a spot across the scene. This allows us to use a high temporal bandwidth device such as an APD to sample the sinusoidal variations in the backscattered object light. This capability, however, comes at a price. A full 3D scan of 100×2500 points requires a total acquisition time of 5 seconds. The obvious shortcoming of raster scanning is the difficulty in handling dynamic scenes. From an engineering standpoint, the problem may be alleviated to an extent by using fast scanning mirrors. These are a few of the many possibilities that are currently being examined. In future work, we hope to investigate the possibility of developing a snapshot-based SH-ToF system that alleviates the difficulties in a sequential point scanning architecture.

6. Conclusion

To summarize, we propose a framework of SH-ToF imaging, which provides micro resolution at a macro imaging range. Our SH-ToF provides much better depth resolution than current commercially available ToF cameras, bridging the gap in precision and sensitivity between optical interferometry and state-of-the-art ToF sensors. A particularly attractive feature of our approach is the flexibility in the choice of depth resolution and unambiguous measurement range. We have demonstrated a depth precision as high as 9 microns for point scanning, and 100 microns for 3D surface scanning. Future work will focus on rapid multi frequency acquisition to enable dynamic high-resolution 3D scanning over large 3D volumes.

Acknowledgement

We thank Joshua Yablon for valuable discussions and Anonymous Reviewer for valuable suggestions. All authors were supported, in part, by the DARPA REVEAL program. M.G. was partially supported by ONR N000141612995. O.C. was partially supported by the NSF CAREER award IIS-145319 and ONR award N00014-15-1-2735.

References

[1] M. Born and E. Wolf. *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Elsevier, 2013.

[2] Y.-Y. Cheng and J. C. Wyant. Two-wavelength phase shifting interferometry. *Appl. Opt.*, 23(24):4539–4543, 1984.

[3] Y.-Y. Cheng and J. C. Wyant. Multiple-wavelength phase-shifting interferometry. *Appl. Opt.*, 24(6):804–807, 1985.

[4] CloudCompare. *[GPL software] (2017), Retrieved from http://www.cloudcompare.org/*.

[5] R. Dändliker, R. Thalmann, and D. Prongué. Two-wavelength laser interferometry using superheterodyne detection. *Opt. Lett.*, 13(5):339–341, 1988.

[6] A. F. Fercher, H. Z. Hu, and U. Vry. Rough surface interferometry with a two-wavelength heterodyne speckle interferometer. *Appl. Opt.*, 24(14):2181–2188, Jul 1985.

[7] S. L. Floch, Y. Salvadé, N. Droz, R. Mitouassiwou, and P. Favre. Superheterodyne configuration for two-wavelength interferometry applied to absolute distance measurement. *Appl. Opt.*, 49(4):714–717, Feb 2010.

[8] S. Foix, G. Alenya, and C. Torras. Lock-in time-of-flight (tof) cameras: a survey. *IEEE Sens. J.*, 11(9), 2011.

[9] C. G. Gordon. Generic vibration criteria for vibration-sensitive equipment. In *Proc. SPIE*, volume 3786, pages 22–39, 1999.

[10] M. Gupta, S. K. Nayar, M. B. Hullin, and J. Martin. Phasor imaging: A generalization of correlation-based time-of-flight imaging. *ACM Trans. Graph.*, 34(5):156, 2015.

[11] G. Häusler. Speckle and Coherence, Encyclopedia of Modern Optics, Elsevier, Academic Press., Oxford, Vol. 1, 114 -123 (2004).

[12] Y. He, B. Liang, J. Yang, S. Li, and J. He. An iterative closest points algorithm for registration of 3D laser scanner point clouds with geometric features. *Sensors*, 17(8):1862, 2017.

[13] F. Heide, M. B. Hullin, J. Gregson, and W. Heidrich. Low-budget transient imaging using photonic mixer devices. *ACM Trans. Graph.*, 32(4):45, 2013.

[14] A. Kadambi and R. Raskar. Rethinking machine vision time of flight with ghz heterodyning. *IEEE Access*, PP(99):1–1, 2017.

[15] A. Kadambi, J. Schiel, and R. Raskar. Macroscopic interferometry: Rethinking depth estimation with frequency-domain time-of-flight. In *Proc. CVPR*, pages 893–902, 2016.

[16] A. Kadambi, V. Taamazyan, B. Shi, and R. Raskar. Polarized 3d: High-quality depth sensing with polarization cues. In *Proc. ICCV*, pages 3370–3378, 2015.

[17] A. Kadambi, R. Whyte, A. Bhandari, L. Streeter, C. Barsi, A. Dorrington, and R. Raskar. Coded time of flight cameras: sparse deconvolution to address multipath interference and recover time profiles. *ACM Trans. Graph.*, 32(6):167, 2013.

[18] KinectSensor. (the second version) 2017. *Retrieved from https://developer.microsoft.com/en-us/windows/kinect*.

[19] R. Lange and P. Seitz. Solid-state time-of-flight range camera. *IEEE J. Quantum Electron.*, 37(3):390–397, 2001.

[20] F. Li, H. Chen, A. Pediredla, C. Yeh, K. He, A. Veeraraghavan, and O. Cossairt. CS-ToF: High-resolution compressive time-of-flight imaging. *Opt. Express*, 25(25):31096–31110, Dec 2017.

[21] F. Li, T. Xu, D.-H. T. Nguyen, X. Huang, C. S. Chen, and C. Zhou. Label-free evaluation of angiogenic sprouting in microengineered devices using ultrahigh-resolution optical coherence microscopy. *J. Biomed. Opt*, 19(1):016006–016006, 2014.

[22] F. Li, J. Yablon, A. Velten, M. Gupta, and O. Cossairt. High-depth-resolution range imaging with multiple-wavelength superheterodyne interferometry using 1550-nm lasers. *Appl. Opt.*, 56(31):H51–H56, 2017.

[23] A. A. Michelson. The relative motion of the earth and of the luminiferous ether. *Am. J. Sci.*, (128):120–129, 1881.

[24] M. O'Toole, F. Heide, L. Xiao, M. B. Hullin, W. Heidrich, and K. N. Kutulakos. Temporal frequency probing for 5d transient analysis of global light transport. *ACM Trans. Graph.*, 33(4):87, 2014.

[25] Y. Salvadé, N. Schuhler, S. Lévêque, and S. Le Floch. High-accuracy absolute distance measurement using frequency comb referenced multiwavelength source. *Appl. Opt.*, 47(14):2715–2720, 2008.

[26] R. Schwarte, Z. Xu, H.-G. Heinol, J. Olk, R. Klein, B. Buxbaum, H. Fischer, and J. Schulte. New electro-optical mixing and correlating sensor: facilities and applications of the photonic mixer device (pmd). In *Proc. SPIE*, volume 3100, pages 245–254, 1997.

[27] S. Shrestha, F. Heide, W. Heidrich, and G. Wetzstein. Computational imaging with multi-camera time-of-flight systems. *ACM Trans. Graph.*, 35(4):33, 2016.

[28] ToF-Sensor. [Texas Instruments], 2017. *Retrieved from http://www.ti.com/tool/OPT8241-CDK-EVM*.

[29] L. Xiao, F. Heide, M. O'Toole, A. Kolb, M. B. Hullin, K. Kutulakos, and W. Heidrich. Defocus deblurring and super-resolution for time-of-flight depth cameras. In *Proc. CVPR*, pages 2376–2384, 2015.

[30] J. Xie, R. S. Feris, and M.-T. Sun. Edge-guided single depth image super resolution. *IEEE Trans. Image Process.*, 25(1):428–438, 2016.