SH-ToF:
Micro Resolution Time-of-Flight Imaging with Superheterodyne Interferometry (Supplementary Materials)

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

Abstract

The current document serves as a companion to the manuscript titled 'SH-ToF: Micro Resolution Time-of-Flight Imaging with Superheterodyne Interferometry'. It is intended to provide details on select topics that were omitted in main manuscript. Section 1 provides detailed mathematical model behind the proposed SH-ToF imager. Section 2 examines geometric distortions in the estimated depth, arising from the use of a 2-axis galvo mirror in our prototype SH-ToF imager. Section 3 describes the geometric distortion compensation technique adopted in the main manuscript for correcting the 3D scanning reconstruction. Section 4 is devoted to an examination of the speckle noise affecting our depth measurements.

1. Mathematical model for SH-ToF

In this section, we provide detailed explanation of mathematical model behind SH-ToF which refers to Section 3.1. First, we can identify the expressions for the complex-valued instantaneous field strength to APD from the sample and reference arms, which are disclosed below [1, 3]:

$$\begin{aligned} E_s(t) = \alpha_1\beta E_1 \exp\{i(2\pi\nu_1 t + \frac{Vt}{\lambda_1} + \frac{4\pi}{\lambda_1}L_s)\} \\ + \alpha_2\beta E_2 \exp\{i(2\pi\nu_2 t + \frac{Vt}{\lambda_2} + \frac{4\pi}{\lambda_2}L_s)\} \end{aligned} \tag{1}$$

$$\begin{aligned} E_r(t) = \gamma_1 E_1 \exp\{i(2\pi\nu_1 t + 2\pi f_{m1} t + \frac{4\pi}{\lambda_1}L_r)\} \\ + \gamma_2 E_2 \exp\{i(2\pi\nu_2 t + 2\pi f_{m2} t + \frac{4\pi}{\lambda_2}L_r)\} \end{aligned} \tag{2}$$

where the terms $E_{1,2}$ represent the field strength of the sample beams at the two wavelengths $\lambda_{1,2}$ respectively. The terms $\alpha_{1,2}$ and $\gamma_{1,2}$ represent losses due to beam splitting and fiber insertion loss. The term β represents the object albedo, while V represents the vibration velocity.

The terms $L_{s,r}$ represent the physical path lengths of the sample and reference arms. The term L_s 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.

As stated previously, the backscattered field E_s from the object inteferes independently with the reference field E_r for the two wavelengths $\lambda_{1,2}$. Following square-law detection, the instantaneous irradiance at the APD may be obtained as the squared modulus of this interference pattern. The expression for the instantaneous irradiance is

$$\begin{aligned} I(t) &= (E_s(t) + E_r(t)) \times (E_s^*(t) + E_r^*(t)) \\ &= \sum_{n=1}^{2} (E_{sn}(t) + E_{rn}(t)) \times (E_{sn}^*(t) + E_{rn}^*(t)) \\ &= a_0 + a_1 \cos(\frac{Vt}{\lambda_1} + \frac{4\pi L}{\lambda_1} - 2\pi f_{m1} t) \\ &+ a_2 \cos(\frac{Vt}{\lambda_2} + \frac{4\pi L}{\lambda_2} - 2\pi f_{m2} t) \end{aligned} \tag{3}$$

where $a_0 = (\alpha_1\beta E_1)^2 + (\gamma_1 E_1)^2 + (\alpha_2\beta E_2)^2 + (\gamma_2 E_2)^2$, $a_1 = 2\alpha_1\beta\gamma_1 E_1^2$, and $a_2 = 2\alpha_2\beta\gamma_2 E_2^2$ are scalar constants. n refers to the n^{th} laser.

From Eq. 3, it 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 $\frac{4\pi L}{\lambda_{1,2}} = \frac{4\pi(L_s - L_r)}{\lambda_{1,2}}$ 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}. The APD output is then digitized using a high-speed DAQ whose sampling rate exceeds $2 \times \max(f_{m1}, f_{m2})$. A squaring operation $S(t)$ applied to the digitized signal yields a time varying signal as disclosed below.

$$\begin{aligned} S(t) = [a_0 + a_1 \cos(\frac{Vt}{\lambda_1} + \frac{4\pi L}{\lambda_1} - 2\pi f_{m1} t) \\ + a_2 \cos(\frac{Vt}{\lambda_2} + \frac{4\pi L}{\lambda_2} - 2\pi f_{m2} t)]^2 \end{aligned} \tag{4}$$

The signal $S(t)$ contains multiple frequency components at f_{m1}, f_{m2}, $(f_{m1} - f_{m2})$, $(f_{m1} + f_{m2})$, $2f_{m1}$, and $2f_{m2}$

Hz as below:

$$
\begin{aligned}
S(t) = & \frac{2a_0^2 + a_1^2 + a_2^2}{2} + 2a_0a_1\cos\left(\frac{Vt}{\lambda_1} + \frac{4\pi L}{\lambda_1} - 2\pi f_{m1}t\right) \\
& + 2a_0a_2\cos\left(\frac{Vt}{\lambda_2} + \frac{4\pi L}{\lambda_2} - 2\pi f_{m2}t\right) \\
& + a_1a_2\cos\left(\frac{Vt}{\lambda_1} + \frac{Vt}{\lambda_2} + \frac{4\pi L}{\lambda_1} + \frac{4\pi L}{\lambda_2} - 2\pi(f_{m1} + f_{m2})t\right) \\
& + a_1a_2\cos\left(\frac{Vt}{\lambda_1} - \frac{Vt}{\lambda_2} + \frac{4\pi L}{\lambda_1} - \frac{4\pi L}{\lambda_2} - 2\pi(f_{m1} - f_{m2})t\right) \\
& + \frac{a_1^2}{2}\cos\left[2\cdot\left(\frac{Vt}{\lambda_1} + \frac{4\pi L}{\lambda_1} - 2\pi f_{m1}t\right)\right] \\
& + \frac{a_2^2}{2}\cos\left[2\cdot\left(\frac{Vt}{\lambda_2} + \frac{4\pi L}{\lambda_2} - 2\pi f_{m2}t\right)\right]
\end{aligned}
$$

A band-pass filter is then used to pick off the beat-note frequency component with frequency of $(f_{m1} - f_{m2})$, and the rest mathematical model follows as that described in the main manuscript.

2. Ground truth depth estimation for line scan on planar surface

The current section references the experiment described in Section 4.2 of the main manuscript. In the experiment, it was observed that the depth profile associated with a strictly planar object exhibits curvature. The curvature stems from the use of a 2-axis galvo mirrors to scan/image object points. The locus of points traced by the galvo mirrors resides on a hemi-spherical surface. The depth measurements reported by our prototype represent the euclidean distance to the center of this sphere (labelled **O** in Fig. 1). As a result, the raw depth profile of a planar surface appears curved. The artifact is reminiscent of geometric distortion in traditional imaging.

In subsequent paragraphs, we outline a procedure to characterize this geometric distortion.

The illustration of Fig. 1 will be used to introduce relevant concepts. Suppose the point '**A**' represents the position of the illuminated spot when zero voltage is applied to the galvo mirrors. The line segment **OA** represents the Z-axis of a world coordinate frame centered at the point '**O**'. The depth associated with the point '**A**' (say d) is chosen as the reference depth for recovering the true depth of other points on the planar object. Applying a a sinusoidal voltage to one of the galvo mirrors (fast-axis) causes the illuminated spot to trace a horizontal line on the planar object. Suppose '**B**' represents a point on this line segment. The depth associated with the point '**B**' as reported by our SH-ToF imager is given by:

$$
l = d \cdot \left(\frac{1}{\cos\alpha} + \tan\alpha \cdot \frac{\sin\theta}{\cos(\alpha + \theta)}\right) \tag{5a}
$$

$$
\alpha = V_0 \cdot \sin(2\pi f t) \cdot \gamma \tag{5b}
$$

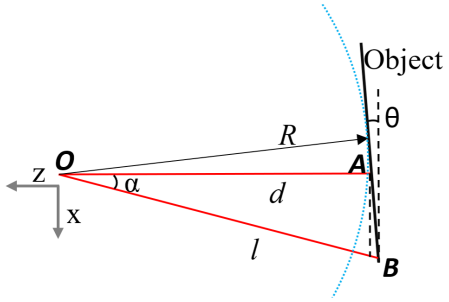

Figure 1: **Ground truth depth estimation for planar surface:** The beam is scanning along a line on the object. O: the scanning center of the Galvo.

where V_0: the amplitude of the sinusoid controlling signal on the Galvo. f: the frequency of the controlling signal on Galvo. γ: the scale factor for the Galvo. The angle θ represents the angle made by the planar facet with respect to the XY plane of the world coordinate frame.

A series of distance measurements with SH-ToF imager $\{l_m\}$ are recorded as the illuminated spot traces a horizontal line, by sweeping the fast-axis mirror through the angles $\{\alpha\}$. Knowledge of d, the mirror angles $\{\alpha\}$ and the associated distance measurements $\{l_m\}$ are used to solve for the tilt angle θ of the planar object.

The fitted ground truth depth profiles $\{l\}$ (solid red plots) in the line-scan measurements of Fig. 7(f-j)) within the main manuscript were obtained by estimating θ from the noisy measurements (in blue), and then plotting Eq.(1a) for known values of d and mirror angles $\{\alpha\}$.

The procedure outlined thus far can be used to identify the geometric distortion in the horizontal direction. The next section outlines a procedure for identifying the distortion in the vertical direction.

3. Depth compensation for 3D scanning

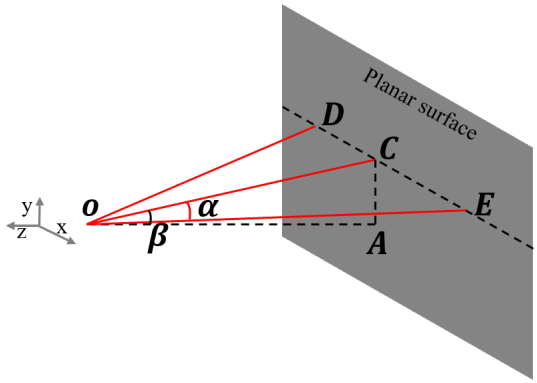

Figure 2: **Depth compensation for optical geometry distortion in 3D scanning: O**: scanning center on the Galvo.

This section references the experiments described in

Section 4.3 of the main manuscript. It describes the depth compensation due to the optical geometry when performing 3D scanning. The illustration of Fig. 2 will be used to introduce relevant concepts. The point '**A**' represents the position of the illuminated spot when zero voltage is applied to the galvo mirrors. Applying a linear voltage ramp to the second galvo mirror (slow-axis) causes the illuminated spot to trace a vertical line on the planar object. Suppose '**C**' is a point on this vertical line segment. Applying a sinusoidal voltage to the fast-axis of the galvo mirrors causes the illuminated spot to move horizontally tracing the line segment **DCE**. The depth associated with the point '**E**' as reported by our SH-ToF imager is given by:

$$OE = OC \cdot \left(\frac{1}{\cos\alpha} + \tan\alpha \cdot \frac{\sin\theta}{\cos\left(\alpha+\theta\right)} \right) \tag{6a}$$

$$OC = OA/\cos\beta \tag{6b}$$

$$\alpha = V_0 \cdot \sin(2\pi f t) \cdot \gamma \tag{6c}$$

$$\beta = V_y \cdot \gamma \tag{6d}$$

where α is the angle Galvo mirror moved in x-axis (fast axis). β is the angle Galvo moved in y-axis (slow axis). θ is the planar surface tilted angle. V_y: voltage added on y-axis mirror.

Given a series of distance measurements along the line segment **DCE**, we can solve for the tilt angle of the planar facet compareed to the XZ plane, and also identify the geometric distortion in our depth measurements as seen in the vertical direction.

Knowledge of the geometric distortion in the horizontal and vertical directions is used to compensate the raw depth values reported by the SH-ToF prototype and recover the correct depth.

4. Noise

The final section of this document deals with the sources of noise that affect the depth values reported by our prototpye SH-ToF imager, which refers to Section 5.1 in the main manuscript.

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. In subsequent paragraphs, we examine each noise source in detail.

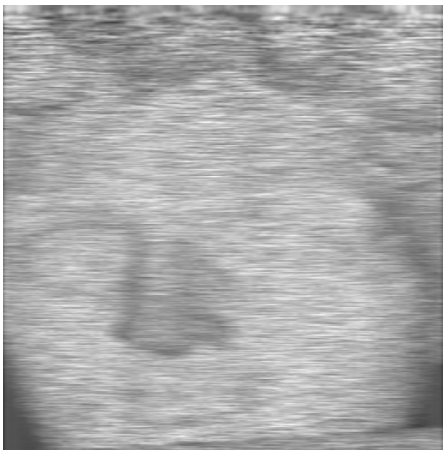

Figure 3: Intensity reconstruction with SH-ToF.

4.1. Speckle Noise

Real world objects and surfaces will likely exhibit microscopic variations in height that are random and comparable to the wavelength of the laser sources used in the SH-ToF imager. Such variations impart a stochastic structure (speckle) to our measurement, due in large part to the randomized interference induced by the height fluctuations. The net result is increased measurement uncertainty in the depth estimated by our ToF imager. The impact of speckle is apparent in the intensity image of the object recovered by our ToF imager. Intuitively, one would expect the image of an object with uniform albedo (such as the bust of David used in our 3D scanning experiment) to exhibit little variation in intensity. Inspection of the albedo image in Fig.3 (reconstruction with Eq.7) confirms the presence of random intensity variation (or speckle) arising from microscopic roughness of the object. Image filtering techniques could be used to mitigate the impact of speckle noise in the albedo image.

The impact of speckle on the estimated depth is a bit more complicated. Firstly, the phase of the backscattered optical field emerging from the object is randomized on account of speckle. As a result, the phase terms ($4\pi L/c\nu_1$,$4\pi L/c\nu_2$) in the expression for the APD irradiance (Eq. 3) is randomized. The issue is addressed by using two closely spaced frequencies (ν_1, ν_2), as in our approach. It is observed that the phase fluctuations in the speckle patterns associated with the two wavelengths are largely identical so that the phase difference $4\pi L/c(\nu_1 - \nu_2)$ is nearly immune to the phase noise arising from surface roughness.

Secondly, the phase of the optical field in the vicinity of a speckle minimum (dark spot for which $a_1 \approx 0$ and/or $a_2 \approx 0$ in Eq. 3 is unreliable. As a result, the numerical estimate of the phase difference $\Phi(L)$ in Eq. 6, and consequently the depth is also unreliable. This suggests that the fidelity of a depth measurement is largely determined by the signal strength observed at the APD. The above observation

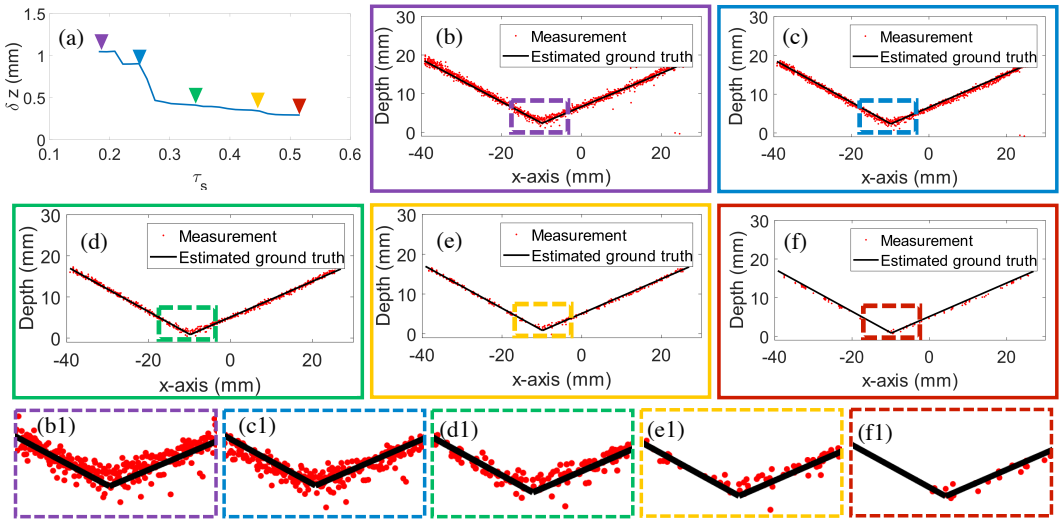

Figure 4: **Depth precision with different threshold (τ_s) on the amplitude of signal S(t):** (a). Depth precision (δ_z) with different threshold (τ_s) values. (b-f). Depth estimations along x axis and the estimated ground truth with threshold values marked with corresponding colorful triangles in (a). Close-up images (b1-f1) show more details of the 'V' shape.

serves as the basis of a strategy for discarding unreliable measurements by thresholding the signal from the APD. The experimental results disclosed in section 4.2 and 4.3 are the outcome of thresholding the APD signal.

Figure 4 demonstrates the benefit of thresholding for one line scan of the folded cardboard (Fig. 9(b)) described in section 4.3.1. It is evident that aggressive thresholding provides increased noise immunity, albeit at the expense of point-cloud density.

In summary, the fidelity and precision of our depth measurements are largely limited by speckle noise and not photon noise as is the case with traditional ToF cameras.

4.2. Laser Emission Frequency Drift (Section 4.3.2)

An advantage of the proposed SH-ToF framework is the ability to tradeoff the unambiguous measurement range for depth resolution. This is made possible by tuning the emission wavelength of the laser sources in our prototype. The tuning is realized by adjusting the temperature of the laser driver module. Unfortunately, the tunability comes at the expense of fluctuations in the emission wavelength/frequency. It is observed that the laser emission frequency drifts by as much as 20 MHz over several seconds. This drift is unlikely to affect a single depth measurement as the drift time scale far exceeds the acquisition time of a single measurement. However, the laser frequency drift will likely impact the full-field 3D scanning measurements.

As a result of this drift, the frequency of the time-varying sinusoid encoding scene depth exhibits undesired temporal variations. The behavior is evident in the expression for the phase associated with a measurement at the synthetic wavelength, disclosed below

$$\Phi = \text{mod}\left(\frac{4\pi[(\nu_1 - \nu_2) + \delta\nu]L}{c}, 2\pi\right) \tag{7a}$$

$$L = \Phi \cdot \Lambda/(4 \cdot \pi) \tag{7b}$$

The term $\delta\nu$ represents the drift in the synthetic frequency of the Superheterodyne Interferometer.

Fortunately, the issue of laser frequency drift may be addressed using approaches such as offset phase locking [2]. Alternatively, we could employ a reference planar surface adjacent to the object, and scan over the object and the reference surface. The planarity of the reference surface may be exploited in aligning the depth-measurements from each raster-scan of the galvo mirror. A third alternative involves the use of a second 'reference arm' that solely monitors the drift in the synthetic frequency. This knowledge could then be used to compensate for the effect of frequency drift in our depth measurements.

4.3. Vibration

It is common knowledge that interferometers are sensitive to variations in the path of the interferometer arms. Such variations are induced by enviromental fluctuations including air currents and mechanical vibrations. The problem is typically mitigated by mounting the interferometer on a vibration isolation platform such as an optical table, while enclosing the entire structure. The above approach is impratical for applications such as 3D scanning.

Fortunately, the Superheterodyne Interferometer is only sensitive to vibrations that are in the order of the synthetic wavelength, which is typically of the order of millimeters. Furthermore, the light path from the galvo to the scene in our ToF imager is common to both arms of the Supehetero-

dyne Interferometer. As a result, it suffices to mount just the interferometer on a stable mechanical platform such as an optical breadboard equipped with vibration dampers.

References

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

[2] F. Friederich, G. Schuricht, A. Deninger, F. Lison, G. Spickermann, P. H. Bolívar, and H. G. Roskos. Phase-locking of the beat signal of two distributed-feedback diode lasers to oscillators working in the mhz to thz range. *Opt. Express*, 18(8):8621–8629, 2010.

[3] 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.