



Open architecture time of flight 3D SWIR camera operating at 150 MHz modulation frequency

SKYLER WHEATON,^{1,2} ALIREZA BONAKDAR,^{1,2} IMAN HASSANI NIA,¹ CHEE LEONG TAN,¹ VALA FATHIPOUR,¹ AND HOOMAN MOHSENI^{1,*}

¹Bio-Inspired Sensors and Optoelectronics Laboratory, Northwestern University, Evanston, IL, USA

²Authors contributed equally to this work

*hmohseni@northwestern.edu

Abstract: In the past two decades 3-D cameras have proven to be one of the next revolutions in machine vision. However, these devices are still an emerging technology with a particularly narrow set of commercially available devices. In this paper, the concept and execution of the first short wavelength infrared (SWIR) time-of-flight (ToF) 3-D camera system operating at a wavelength of 1550 nm is presented. By decoupling the optical and electrical components of the system in an open architecture we not only surpass many of the limitations of an on-chip integrated solution, but also can easily change the imaging device based on the requirements of the application. We achieve modulation frequencies up to 150 MHz, which exceeds the conventional values currently published for other large format modulators by about five times. This increase in the modulation frequency allows for a TOF camera with significantly higher depth resolution, while the open architecture design allows for a highly reconfigurable device that can be modified for specific working conditions.

©2017 Optical Society of America

OCIS codes: (110.6880) Three-dimensional image acquisition; (150.6910) Three-dimensional sensing; (110.0110) Imaging systems

References and links

1. C. Niclass, A. Rochas, P. A. Besse, and E. Charbon, "Design and characterization of a CMOS 3-D image sensor based on single photon avalanche diodes," *IEEE J. Solid-State Circuits* **40**(9), 1847–1854 (2005).
2. O. Elkhaili, O. M. Schrey, P. Mengel, M. Petermann, W. Brockherde, and B. J. Hosticka, "A 4x64 pixel CMOS image sensor for 3-D measurement applications," *IEEE J. Solid-State Circuits* **39**(7), 1208–1212 (2004).
3. R. Lange and P. Seitz, "Solid-state time-of-flight range camera," *IEEE J. Quantum Electron.* **37**(3), 390–397 (2001).
4. T. Oggier, M. Lehmann, R. Kaufmann, M. Schweizer, M. Richter, P. Metzler, G. Lang, F. Lustenberger and N. Blanc, "An all-solid-state optical range camera for 3D real-time imaging with sub-centimeter depth resolution," *Proc. SPIE Optical Design and Engineering* **5249**, 534 (2004).
5. M. Frank, M. Plaue, H. Rapp, U. Kothe, B. Jahne, and F. Hamprecht, "Theoretical and experimental error analysis of continuous-wave time-of-flight range cameras," *Opt. Eng.* **48**, 013602 (2009).
6. T.-Y. Lee, D. Min, S. Lee, W. Kim, J. Jung, I. Ovsiannikov, Y. Jin, Y. Park, and E. Fossum, "A Time-of-Flight 3-D image sensor with concentric-photogates demodulation pixels," *IEEE Trans. Electron Dev.* **61**(3), 870–877 (2014).
7. H. Mohseni, W. K. Chan, H. An, A. Ulmer, and D. Capewell, "High-performance surface-normal modulators based on stepped quantum wells," *SPIE* **5814**, 191–198 (2005).
8. Y. Park, Y. Cho, J. You, C. Park, H. Yoon, S. Lee, J. Kwon, S. Lee, B. Hoon, Y. Lee, "Three-dimensional imaging using fast micromachined electro-absorptive shutter," *J. Micro./Nanolith. MEMS MOEMS*, **12**(2) 023011 (2013).
9. M. Kawakita, K. Iizuka, T. Aida, H. Kikuchi, H. Fujikake, J. Yonai, and K. Takizawa, "Axi-Vision Camera (real-time distance-mapping camera)," *Appl. Opt.* **39**(22), 3931–3939 (2000).
10. H. Horvath, "Atmospheric light absorption—A review," *Atmos. Environ. Part A. Gen. Topics* **27**(3), 293–317 (1993).
11. D. H. Sliney, and J. Mellerio, *Safety with Lasers and Other Optical Sources: A Comprehensive Handbook*, 88, 273, (Springer, 1980).
12. C. S. Bamji, P. O'Connor, T. Elkhatib, S. Mehta, B. Thompson, L. A. Prather, D. Snow, O. C. Akkaya, A. Daniel, A. D. Payne, T. Perry, M. Fenton, and V. Chan, "A 0.13 μm CMOS System-on-Chip for a 512x 424

Time-of-Flight image sensor with multi-frequency photo-demodulation up to 130 MHz and 2 GS/s ADC," IEEE J. Solid-State Circuits **50**(1), 303–319 (2015).

1. Introduction

The capability of the humans to perceive the world in three dimensions is a key aspect in the development of our species. While nature has handled this for us, accurate 3-D imaging has presented itself as a major hurdle in man-made machinery [1, 2]. Modern solutions that seek to surpass this hurdle find themselves facing very stringent design constraints that can heavily restrict the performance of 3D imaging systems. Of particular interest in the last few years has been the concept of CW time of flight (ToF) cameras. These devices use an array of pixels that concurrently measure the depth of different regions of the field of view (FoV), which is broadly illuminated with a modulated optical signal [1]. The depth information is extracted by demodulating the reflected light off of the scene to unfold the phase of the returned signal [3]. These systems operate in a regime where high frame rate is achievable. However, this is typically at the cost of depth accuracy and is limited by the illumination modulation / demodulation frequency and the background signal level [4]. While these devices are a significant step forward for 3-D imaging, they suffer from their reliance on fast on-chip electronics to achieve high modulation frequencies. Currently, the best 3D imaging ToF devices on the market are based on photon mixing devices (PMD) and operate in the range of 20 - 130 MHz modulation frequency [5, 6]. This is particularly important for CW ToF systems, due to the depth resolution linearly scaling with the modulation frequency.

In this paper, we demonstrate the first ToF SWIR camera operating at 150 MHz modulation frequency. A key innovation in this system in comparison to a PMD device is the open architecture configuration, which uses an efficient large format electro-absorptive modulator [7] as a global shutter to demodulate the reflected light from the scene before it is captured by an operationally independent image sensor. This configuration enables a free selection of the imaging chip, which can be non-silicon based. A 30 MHz TOF camera operating with a large format modulator acting as a global shutter was demonstrated in the wavelength range of 850-980 nm by Park et al in [8]. However, a theoretical model and comparisons to such a model were not presented. Similar methods that make use of image intensifiers as a type of shutter have been demonstrated in the visible wavelengths [9]. Here we demonstrate operation in the SWIR band using a similar methodology that we show is related very closely to the operation of a PMD sensor by comparing our measured performance to an accepted theory of operation for PMD devices. The appeal of a SWIR based system stems from the absorption of the SWIR band of solar radiation by our atmosphere at 1530 - 1550 nm [10]. By exploiting this property, SWIR based 3-D imagers have the potential to work in outdoor spaces while still maintaining a high depth resolution due to a strong reduction of parasitic background illumination. Additionally, it has been shown that wavelengths beyond 1.4 microns cannot effectively focus onto the retina due to the strong optical absorption by the water content of the eye [11]. This makes the SWIR band an ideal choice for applications involving human users and bystanders considering the maximum eye-safe power in the SWIR band is 100 - 1000 times larger than the maximum power in the Near IR (NIR) band in the range of 850 – 980 nm [10].

2. 4-Phase-shifting technique

The 3D camera system is comprised of three main components, an illumination source an optical modulator and a camera. The illumination source is modulated at a fixed frequency, the light emitted is incident on a scene and is recollected after passing through a stepped quantum well modulator (SQM) which acts as a fast shutter. The SQM is modulated at the same base frequency but can be phase shifted by either 0, 90, 180, or 270 degrees. By taking an image at each of the four phase shift states, the 3D depth of the scene can be reconstructed

for reach camera pixel based on the measured optical intensity at each state [12]. To form a theory of operation we assume our output optical illumination signal has the following form:

$$I(t) = P_{avg} + P_{amp}[\cos(\omega t)]. \quad (1)$$

where P_{avg} is the offset due to background optical signal, P_{amp} is the amplitude of modulation of the illumination source in units of watts, and ω is the angular modulation frequency. This signal is incident on the scene, the return signal is then of the form:

$$I(t) = P_{avg} + P_{amp}[\cos(\omega t + \rho)]. \quad (2)$$

where ϕ is the additional phase accumulated due to the round-trip path length. In this formulation, the scene reflectivity is assumed to be unity. The demodulation signal of the modulator can be similarly expressed as:

$$T(t) = T_{min} + T_{amp}[1 + \cos(\omega t + \rho)]. \quad (3)$$

where T_{min} is the average normalized minimum transmission through the SQM, T_{amp} is the demodulation signal amplitude, lastly ρ is the phase shift associated with the phase sampling point which is chosen to be either 0, 90, 180, or 270 degrees. The signal which is recorded by the camera will follow the form of Eq. (4).

$$I_R(t) = T(t)I(t). \quad (4)$$

by expanding this equation and making the assumption that the integration time of the camera is significantly larger than the period of the modulated signal the function for the average signal seen by the camera can be expressed as:

$$I_{Ravg}(\rho) = T_{avg}P_{avg} + \left(\frac{P_{amp}T_{amp}}{2}\right)\cos(\phi - \rho). \quad (5)$$

where $T_{avg} = T_{min} + T_{amp}$. By sampling Eq. (5) at 4 different values for ρ namely $\rho_0, \rho_{90}, \rho_{180}, \rho_{270}$ we obtain four independent solutions which can be used to discern the phase of the recollected signal from the scene:

$$I_{Ravg}(\rho_{270}) - I_{Ravg}(\rho_{90}) = -(P_{amp}T_{amp})\sin(\phi). \quad (6)$$

$$I_{Ravg}(\rho_0) - I_{Ravg}(\rho_{180}) = (P_{amp}T_{amp})\cos(\phi). \quad (7)$$

therefore, we see that by dividing Eq. (6) and Eq. (7) we recover the same equation for phase estimation as given in [5]:

$$\text{Arctan}\left(\frac{I_{Ravg}(\rho_{270}) - I_{Ravg}(\rho_{90})}{I_{Ravg}(\rho_0) - I_{Ravg}(\rho_{180})}\right) = -\phi. \quad (8)$$

with knowledge of our returned phase and a known modulation frequency the depth can be deduced using Eq. (9):

$$D = \frac{c}{4\pi f_{mod}}\phi. \quad (9)$$

where c is the speed of light in air and f_{mod} is the modulation frequency of both the modulator and the illumination source. Additionally, the amplitude and offset of the 4-Phase

signal (A and B respectively) can be estimated from the 4-Phase image intensities $A_0, A_{90}, A_{180}, A_{270}$ for each pixel of the image:

$$A = \frac{\sqrt{(A_{270} - A_{90})^2 + (A_0 - A_{180})^2}}{2}. \quad (10)$$

$$B = \frac{(A_0 + A_{90} + A_{180} + A_{270})}{4}. \quad (11)$$

these values are labeled in the measurement schematic shown in Fig. 1(b). With knowledge of the characteristics of the device used to measure the 4-phase images these A and B values can be directly related back to the optical modulation amplitude and the optical signal offset. With these equations, from one set of 4-Phase frames the depth can be calculated (Eq. (8) and (9)), the average intensity can be deduced from the value of B and the depth resolution can be estimated from A providing a robust measurement technique [3]. Such a system could support active illumination running on a feedback loops based on the measured values of A and B.

3. SWIR TOF system

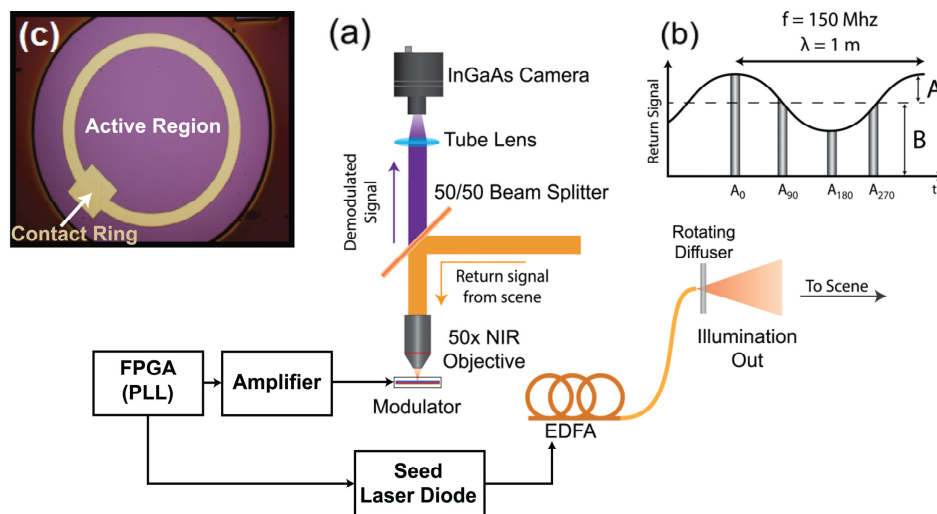


Fig. 1. (a) Schematic view of SWIR TOF experimental setup. (b) Example of reconstructed signal based on the amplitudes of the four phase images, amplitude of the signal (A), and offset (B). (c) Optical microscope image of 1 mm diameter modulator.

The ToF experimental setup can be seen in Fig. 1. After the modulated illumination signal passes through the rotating diffuser to reduce speckle noise, the recollecting signal reflected back from the scene is split by a 50/50 beam splitter, sending half of the received signal into a 50x NIR microscope objective. The objective focuses the received signal onto the back surface of the SQM. The backside of the SQM is coated with gold which acts as a mirror as well as back electrical contact to drive the device. Here the signal is demodulated, reflected, and finally recollecting by the microscope objective. The collimated signal passes through the beam splitter and finally is mapped onto the camera's pixels through a camera lens. The total insertion loss of the system was measured to be 12 dB. The 4-Phase shifted modulation signals are generated by a Spartan VI FPGA board from Xilinx. One of the quadrature signals drives a 1550 nm seed laser for the erbium doped fiber amplifier (EDFA), while the other

drives the modulator through a voltage amplifier (RFBAY MPA12-30). The output of the amplifier is combined with a -50 volt DC bias through a bias-tee.

To characterize the performance of the SQM, a different setup was built that uses a circulator and a collimator to send a collimated beam into the system as opposed to imaging a scene. The CW signal is modulated by the SQM, reflected and then recollimated through the circulator and measured using an InGaAs photodiode (Lumentum ETX 100RFC2) and an oscilloscope. The oscilloscope is used to measure the amplitude of the modulation induced by the SQM which is amplified using a high bandwidth voltage amplifier (DUPVA-1-70). A low bandwidth calibrated power meter is used to measure the DC offset of the modulated signal. With these two values the depth of modulation (DoM) can be defined as:

$$DoM = \frac{2 \text{Amplitude}}{\text{Amplitude} + DC_{offset}}. \quad (12)$$

For our system operating at 150 MHz with a -50 volt DC bias and a 15 volt peak to peak AC signal on the modulator we obtain a DoM of 35%. Analogously the DoM of the illumination source was measured to be 98% leading to a full system contrast which is defined as the illumination DoM multiplied by the modulator (demodulation) DoM of 34.3%

4. Results and discussion

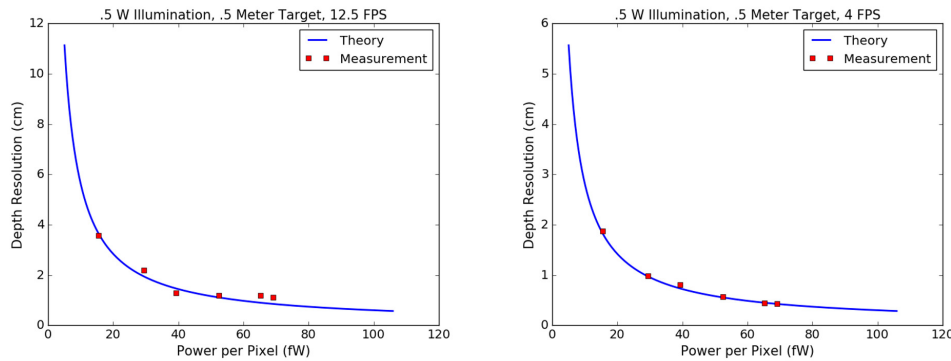


Fig. 2. Depth resolution performance measurement and comparison to theory. (a) Comparison to theory (based on Eq. (13)) when operating at 12.5 fps. (b) Comparison to theory (based on Eq. (13)) when operating at 4.1 fps.

As shown in [3] the depth resolution of a CW ToF system can be estimated with the following equation:

$$\Delta L = \frac{L}{\sqrt{8}} \frac{\sqrt{N_{pseudo} + N_{bg} + PE_{opt}}}{2kPE_{opt}}. \quad (13)$$

where L is the ambiguity range as dictated by the modulation frequency. N_{pseudo} is the summation of the offset due to dark current and read noise, N_{bg} is the counts due to the background signal inherent in the scene, k is the system modulation contrast (product of illumination DoM and demodulation DoM) and PE_{opt} is the returned signal counts. Figure 2(a) shows the comparison of the measured system performance to the theoretical system performance when operating at 12.5 fps and an average illumination power of 0.5 Watts. Figure 2(b) shows the system performance and theory comparison when operating with the same configuration but with increased averaging leading to a reduced frame rate of 4 fps. The background count value for our indoor scene was measured to be 200 photoelectrons per frame. The values measured and used in the model are tabulated below in Table 1.

Table 1. Summary of values used in theoretical model of TOF system.

Variable Name	Value
N_{psuedo}	360000 (e)
N_{bg}	200 (e)
PE_{opt}	6221 – 27800 (e)
k	0.343

To show the ability of our system in capturing minute features within a scene, a 2.5 by 5 inch bust was imaged at various frame rates. The base framerate of our SWIR camera is 50 FPS, thus we are limited to a depth framerate of $50/4 = 12.5$ FPS when taking each of the four-phase images sequentially. We note that if each 4-phase image is taken with a rolling queue-like scheme the depth frame rate can be matched to that of the base camera FPS. Figure 3(a) shows the raw SWIR image output of the camera. The integration time of the camera is 20 ms and the imaging range is 0.6 meters. Figure 3(b-d) show the depth image output at 12.5 FPS, 6 FPS and 1 FPS respectively. The spatially periodic noise seen in the images is an unavoidable pollution of the depth information caused by the noise of our camera's internal charge transimpedance amplifier. Despite this pollution of the image which could be solved by using a camera more suited for high frame rate operation, the depth resolution calculated for the 12.5 FPS image is consistent with theory at a value of 1.35 cm. With additional averaging at the cost of framerate, a depth resolution of 0.65 cm is achievable at 1 FPS readout.

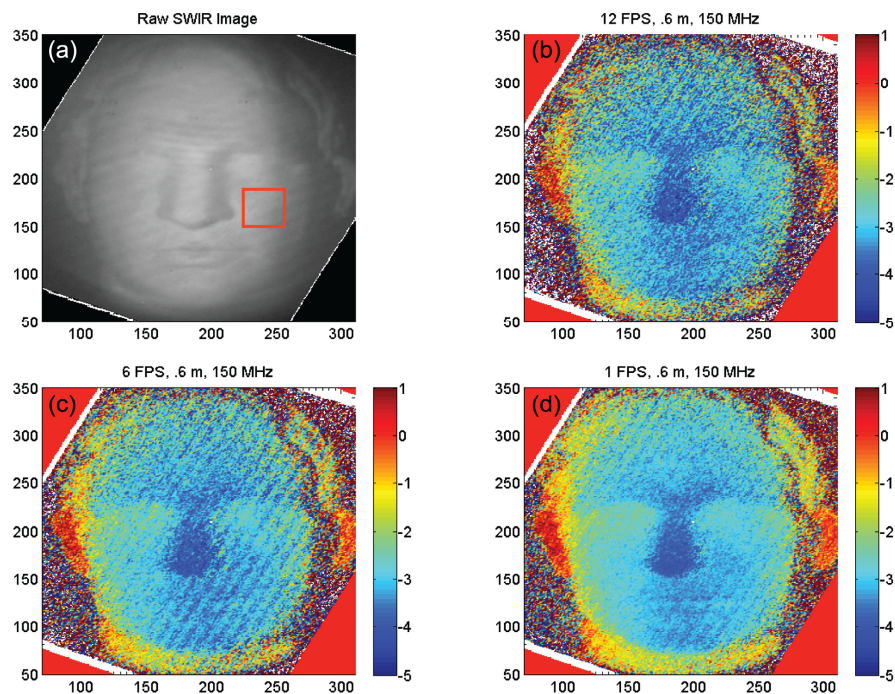


Fig. 3. 3D images of 2.5 by 5-inch statue bust. (a) Raw SWIR image of bust, red subsection shows the region of the image used for the depth resolution calculation at each framerate. (b) Depth output at 12 FPS, resolution = 1.36 cm. (c) Depth output at 6 FPS, resolution = 0.865 cm. (d) Depth output at 1 FPS, resolution = 0.65 cm.

5. Conclusions

In conclusion, we demonstrated time of flight imaging using an open architecture design that alleviates many limitations of the currently available ToF systems. The system presented here operates at a wavelength of 1550 nm in the SWIR band, which hosts many benefits. However, the design can potentially be used with any off the shelf imager which could operate at a different wavelength or be chosen based on noise/performance constraints. The system was shown to operate at modulation frequencies up to 150 MHz, which resulted in depth resolutions on the order of 1.3 - 0.7 cm at a range of 0.6 meters at real time speeds in good agreement with the theory of operation. In the future, we plan to take advantage of our modulator's low capacitance and push for higher modulation frequencies, potentially reaching into the GHz range for a much better depth resolution.

Funding

DARPA/ARO award #W911NF-16-1-0458, ARO award #W911NF-11-1-0390, and NSF award #IIP-1500314 001.

Acknowledgment

We acknowledge the computational resources and staff contributions provided by the Quest high performance computing facility at Northwestern University, which is jointly supported by the Office of the Provost, the Office for Research, and Northwestern University Information Technology.