

Computational Imaging with Multi-Camera Time-of-Flight Systems

Shikhar Shrestha^{1*} Felix Heide^{1,2*} Wolfgang Heidrich^{3,2} Gordon Wetzstein¹
¹Stanford University ²University of British Columbia ³KAUST

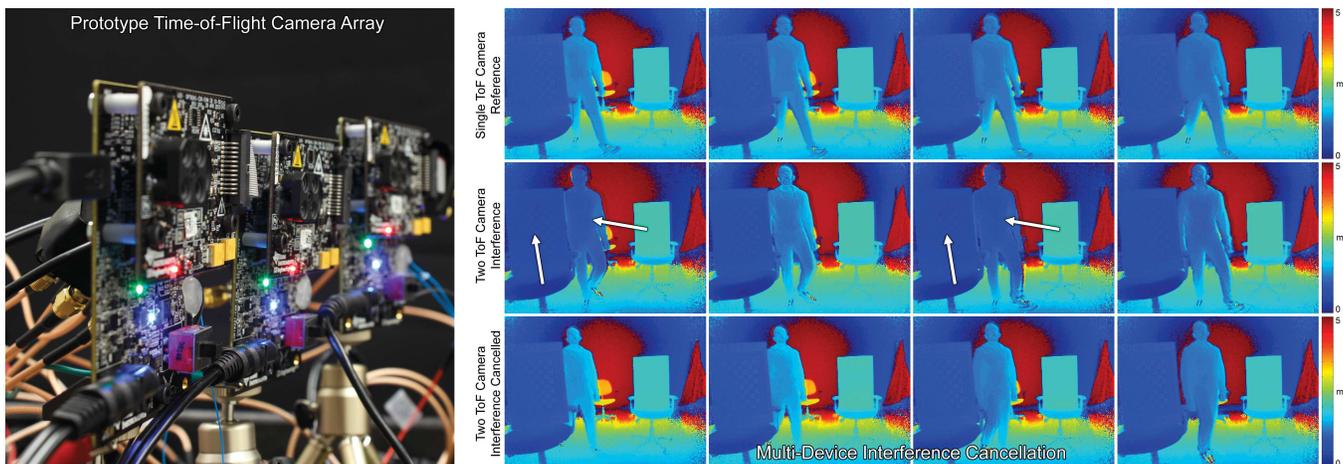


Figure 1: We explore computational imaging with multi-camera time-of-flight systems. Our prototype (left) uses commercially-available sensors, but we design and build external signal generation and control electronics to synchronize the exposures of up to three sensors and drive them with programmable waveforms. One of many applications is multi-device interference cancellation (right). When two time-of-flight cameras are used simultaneously (right), their temporally-modulated illumination codes interfere with one another, which creates periodic artifacts in the estimated depth maps. Operating each light source-camera pair at a different modulation frequency solves this problem. We explore this and other applications of computational multi-camera time-of-flight systems.

Abstract

Depth cameras are a ubiquitous technology used in a wide range of applications, including robotic and machine vision, human-computer interaction, autonomous vehicles as well as augmented and virtual reality. In this paper, we explore the design and applications of phased multi-camera time-of-flight (ToF) systems. We develop a reproducible hardware system that allows for the exposure times and waveforms of up to three cameras to be synchronized. Using this system, we analyze waveform interference between multiple light sources in ToF applications and propose simple solutions to this problem. Building on the concept of orthogonal frequency design, we demonstrate state-of-the-art results for instantaneous radial velocity capture via Doppler time-of-flight imaging and we explore new directions for optically probing global illumination, for example by de-scattering dynamic scenes and by non-line-of-sight motion detection via frequency gating.

Keywords: computational photography, time-of-flight, light fields

Concepts: •Computing methodologies → Computational photography; 3D imaging;

1 Introduction

The emergence of RGB-D or range imaging has had a profound impact on research and technology development. Since the release of the Microsoft Kinect consumer depth camera in 2010, these devices have been widely deployed in living rooms around the world, and they have become valuable tools for many computer vision and

computer graphics researchers, with applications such as scene reconstruction and understanding, pose estimation, action recognition, localization and mapping, navigation, tracking, segmentation, recognition, feature extraction, and reconstruction of geometry, material properties, or lighting conditions (see [Gall et al. 2014] for an overview). Beyond computer vision applications, range imaging is useful for human-computer interaction [Shotton et al. 2011], biometrics, autonomous vehicle and drone navigation, and also for positional tracking of immersive visual computing platforms (augmented and virtual reality, AR/VR). Today, range imaging technology is largely dominated by time-of-flight (ToF) cameras due to their small device form factors, good resolution, robustness in the presence of ambient light, low power, and fast on-chip processing [Hansard et al. 2012].

In this paper, we explore the design and applications of synchronized (i.e. phased) multi-camera systems. While such phased arrays have many interesting applications, they also suffer from potential *multi-device interference* (MDI), that is created when light sources of multiple ToF cameras interact with one another, thereby corrupting the measurements for all sensors. This problem can be understood as a variant of the common *multi-path interference* (MPI) problem that has been well studied for ToF cameras. MPI is an undesirable effect that results in severe degradations of the estimated scene depth; many solutions have been proposed [Fuchs 2010; Dorrington et al. 2011; Jimenez et al. 2012; Kadambi et al. 2013; Freedman et al. 2014; Bhandari et al. 2014; Naik et al. 2015]. Even without global illumination, multi-device interference is a limitation that diminishes the usefulness of ToF cameras for applications where multiple cameras are crucial, such as collaborative work, automotive applications, and AR/VR applications with multiple users. We demonstrate a trivial solution for the MDI problem: these artifacts can be mitigated by driving each sensor/light source pair with sinusoidal waveforms that have different temporal

*Joint first authors

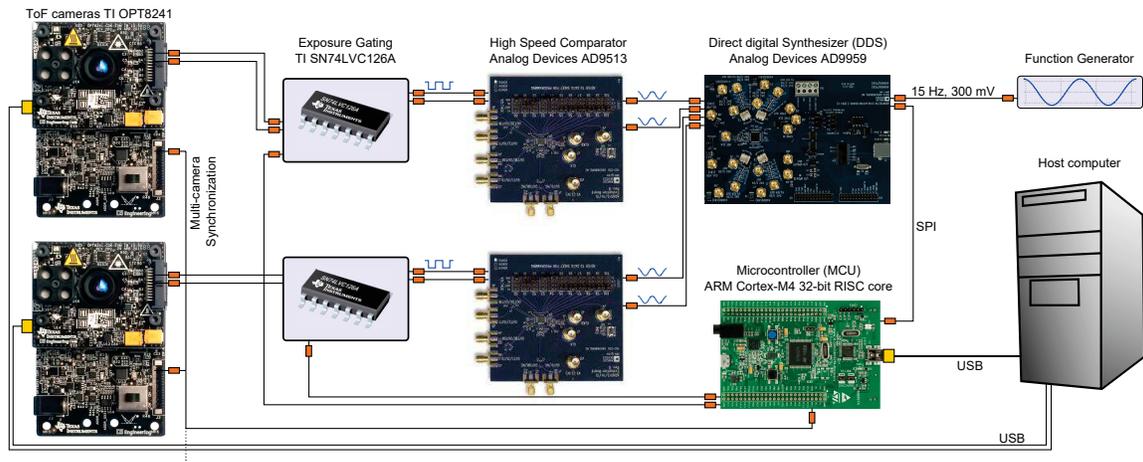


Figure 2: System overview. The proposed multi-camera time-of-flight system is built around the TI OPT8241 camera developer kit (left). All connected cameras are driven in slave mode, such that an external microcontroller (MCU) manages and synchronizes the sensor exposures of all cameras. The MCU also dynamically controls the frequency and relative phase settings of all four direct digital synthesizer (DDS) channels. The DDS generates analog waveforms that are digitized before being fed into the cameras. An additional exposure gating circuit is necessary to guarantee that the cameras receive no residual signals when they are not exposing. Each camera requires two independent waveforms, one for the sensor and another one for the on-board light source.

frequencies.

Computational time-of-flight imaging has proved to be a useful tool for solving many problems in ToF imaging and for facilitating new applications, such as transient imaging [Velten et al. 2013; Heide et al. 2013; Peters et al. 2015], non-line-of-sight imaging [Kirmani et al. 2009; Velten et al. 2012; Heide et al. 2014a], de-scattering [Heide et al. 2014b], light transport analysis [O’Toole et al. 2014], BRDF estimation [Naik et al. 2011], lensless imaging [Wu et al. 2012], and simultaneous range and velocity imaging [Heide et al. 2015]. Many of these approaches build on the idea of waveform design for the temporally-modulated codes of time-of-flight sensors and light sources. We follow this principle, in particular recent proposals on orthogonal frequency design [Heide et al. 2015], but we extend the idea of waveform design to multi-device setups where several sensors and light sources can simultaneously capture and illuminate the scene.

Phased arrays have been employed in synthetic aperture radar for decades (e.g., [McCandless and Jackson 2004]) and are common for sensing systems throughout the electro-magnetic spectrum. Nevertheless, phasing full-field sensors is an emerging research topic. In this paper, we design the first phased multi-camera ToF system and demonstrate unique benefits for instantaneous range and velocity imaging, multi-device interference cancellation, de-scattering of dynamic scenes, non-line-of-sight (NLOS) motion detection, and many other applications.

In particular, we make the following contributions

- We design and build a multi-camera time-of-flight system using consumer devices with external waveform generation and synchronization. Our hardware and software design is publicly available.
- We derive forward and inverse methods for phased ToF imaging with multiple sensors and light sources.
- We demonstrate benefits for a range of applications, including fast, artifact-free range and velocity imaging, multi-device interference cancellation, de-scattering dynamic scenes, and NLOS motion detection via frequency gating.

Limitations All proposed applications benefit from multiple synchronized camera because some or all of the time-sequential phase- or frequency-stepped measurements necessary for most ToF applications could be recorded simultaneously. However, some of the discussion applications (i.e. de-scattering and NLOS motion detection) do not necessarily require multiple cameras. They could, however, benefit from more than one light source, each operated at a different frequency. The proposed hardware system is currently limited to a maximum of four channels, which can operate three sensors and one light source, one sensor and three light sources, or two sensor-light source pairs.

2 Related Work

Time-of-Flight Systems With 512×424 pixels, the Microsoft Kinect for XBOX ONE is the highest-resolution time-of-flight sensor that is currently on the market. Technical details can be found in the publication by Bamji et al. [2015]. Although this sensor can be modulated with up to 130 MHz, only a few discrete frequency settings are used in practice to resolve phase wrapping ambiguities [Payne et al. 2009]. Unfortunately, the firmware controlling on-device signal generation is not accessible and raw sensor data cannot be read out easily. Hence, the Kinect does not provide a viable hardware platform for computational imaging research.

To overcome the limited settings offered by the Kinect sensor, several research groups have proposed custom control electronics designs. For example, Heide et al. [2013] built a programmable time-of-flight camera based on the PMD CamBoard nano system¹. The same design was recently also used to engineer orthogonal heterodyne frequencies for velocity imaging [Heide et al. 2015]. In both cases, a direct digital synthesizer (DDS, see Sec. 4) generates the external waveforms controlling the temporally-coded illumination and sensor demodulation signals. Kadambi et al. [2013] used the same PMD sensor but their signal generation is done with a field-programmable gate array (FPGA), which is more flexible but less precise than a DDS. Unfortunately, the resolution of the PMD sensor is low (160×120 pixels), access to external sensor modula-

¹http://www.pmdtec.com/products.services/reference_design.php

tion is very challenging, synchronization between multiple devices is not supported, and the sensor is discontinued by the manufacturer. Overall, this system is not suitable for building phased camera arrays and it is also not reproducible.

To our knowledge, the epc660 Evaluation Kit offered by ESPOS Photonics² is the most versatile, commercially-available time-of-flight platform. The listed price of \$5,000, however, makes it costly and the system does not offer multi-sensor and multi-light source synchronization. In this paper, we propose a reproducible, fully-programmable time-of-flight camera platform that provides custom waveforms and multi-device synchronization capabilities at low cost. We explore several unprecedented applications offered by multi-camera time-of-flight systems.

Camera and Light Source Arrays Structured [Levoy and Hanrahan 1996; Gortler et al. 1996] and unstructured [Buehler et al. 2001] camera arrays are common for image-based rendering and for cinematic content creation with immersive visual computing platforms [Rander et al. 1997]. For example, the ability to record actors from multiple different viewpoints [Matusik et al. 2000; Carranza et al. 2003] and re-render them from novel perspectives gave rise to new visual effects in the entertainment industry. Growing interest in virtual and augmented reality applications is starting to fuel the development of a diversity of multi-camera systems, which are the primary means to record live-action content for these experiential computing systems. From a computational imaging perspective, multi-camera arrays not only allow for the light field of a scene to be acquired and re-rendered, but they also facilitate spatial and temporal super-resolution, high dynamic range and panoramic imaging, synthetic aperture photography, and a variety of other application [Wilburn et al. 2005]. Similar to multi-camera systems, arrays of light sources have proved useful for many tasks, including surface orientation and shape reconstruction [Woodham 1980] as well as relighting scenes and actors [Debevec et al. 2000].

Most recently, the idea of stereo [Castaneda et al. 2014; Li et al. 2015] and multi-camera [Jayasuriya et al. 2015] time-of-flight setups was proposed. Castaneda et al. [2014] proposed to capture multiple images with different combinations of light sources turned on and off. This approach allows for the contribution of each light source to be isolated, thereby facilitating conventional ToF depth estimation without multi-source interference. In addition, Castaneda et al. showed that the additional phase measurements of multiple light sources can enhance the estimated depth compared to a single ToF camera. Li et al. [2015] demonstrated that the interfering waveforms of multiple ToF cameras distort the estimated depth values of each camera. The authors proposed to average more than 100 frames, which statistically mitigates the estimation error. The work closest to ours is that of Jayasuriya et al. [2015], who proposed a combination of light field and time-of-flight imaging. Their idea was demonstrated by moving a Microsoft Kinect on a translation stage and capturing multiple images sequentially. Applications in synthetic aperture photography and resolving phase wrapping ambiguities make this idea intriguing. Kim et al. [2014] and Ti et al. [2015] recently demonstrated photometric stereo with a single ToF camera and multiple light sources, although they do rely on time-sequential capture for the individual lights.

Although stereo and multi-camera ToF arrays have been shown to provide benefits for depth estimation and other applications, none of the previously-proposed approaches is suitable for dynamic scenes because all of them require multiple images to be captured in sequence. We argue that the full potential of time-of-flight camera arrays is only unlocked by phasing them, that is by synchronizing

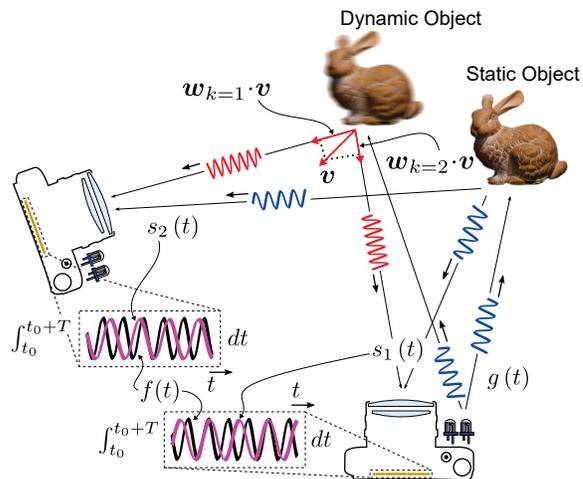


Figure 3: Illustration of phased time-of-flight camera pair. A light source emits a temporally-coded illumination pattern $g(t)$. The light reflected from static objects towards the cameras observes a phase shift (blue signals), whereas the light reflected by dynamic objects is shifted in phase and frequency (red signals). Each sensor correlates the incident signal $s(t)$ with a demodulation signal $f(t)$ that is synchronized among the sensors. The algorithms described in the text recover both depth and 3D velocity of the imaged scene.

their exposures as well as the illumination and sensor demodulation waveforms. In this paper, we design and build such a phased multi-camera time-of-flight system (see Fig. 2). We demonstrate new applications beyond instantaneous range and velocity imaging of dynamic scenes and resolving phase wrapping ambiguities, such as multi-device interference cancellation, de-scattering of objects in motion, and non-line-of-sight motion detection.

Phased cameras were recently also discussed in the context of coherent computational imaging [Dong et al. 2014; Holloway et al. 2015]. These systems exploit interference properties of coherent light to compute images with a resolution beyond the diffraction limit. The continuous-wave time-of-flight principle relies on interference between an amplitude-modulated illumination waveform with the on-sensor demodulation signal. These frequencies are orders of magnitude lower than the wavelength of light. Due to the long wavelengths of the signals used in time-of-flight cameras, the concept of improving image resolution from interference does not apply in the same manner as it does for coherent camera arrays.

3 Phased ToF Array Imaging

3.1 Range Imaging via Homodyne Waveforms

Most time-of-flight cameras operate in amplitude-modulated continuous-wave (AMCW) mode, where a fast, active light source illuminates a scene with a time-varying signal or waveform. Usually, these waveforms are described as sinusoids such that the temporally-varying illumination is

$$g(t) = g_1 \cos(\omega t) + g_0. \quad (1)$$

Here, ω is the AM frequency, which is in the order of 10-130 MHz, and g_0, g_1 are constants. An object at a distance d reflects part of the illumination back to the camera. The following signal is incident on

²<http://www.espos.ch/>

the sensor

$$\begin{aligned} s(t) &= \alpha \cos\left(\omega \left(t - \frac{2d}{c}\right)\right) + b \\ &= \alpha \cos(\omega t + \phi) + b. \end{aligned} \quad (2)$$

This formulation contains an ambient term $b = \alpha g_0 + \beta$, with β being the constant background, as well as the amplitude α , which combines g_1 , the square distance falloff, as well as the albedo of the object. Due to the propagation distance, the phase of the received signal is shifted by $\phi = -2d\omega/c$.

The AMCW time-of-flight problem is an estimation of the phase shift ϕ , and with it, the scene depth. For this purpose, the temporally-varying signal incident on the sensor $s(t)$ is demodulated by a sinusoidal function $f_\psi(t) = \cos(\omega t + \psi)$. In practice, this demodulation is implemented by periodically directing photoelectrons into one of two ‘‘buckets’’ within each sensor pixel (see e.g. [Hansard et al. 2012]). When the illumination and sensor modulation frequencies are the same, the ToF camera operates in *homodyne* mode.

To model the integrated sensor measurements, we account for a finite exposure time T , which acts as a temporal low-pass filter on the demodulated sensor signal. Assuming that $T \gg 1/\omega$, the measured intensity is

$$i_\psi(t') = ((f_\psi(t) s(t)) * \text{rect}_T)(t') \approx \frac{\alpha}{2} \cos(\psi - \phi). \quad (3)$$

The depth phase ϕ is computed from four measurements that step the relative phase difference between illumination and sensor by $\frac{\pi}{2}$ (i.e. $\psi = [0 \frac{\pi}{2} \frac{3\pi}{2} \pi]$) such that

$$\phi_{est} = \tan^{-1}\left(\frac{i_{\frac{3\pi}{2}} - i_{\frac{\pi}{2}}}{i_0 - i_\pi}\right), \text{ and } d_{est} = \frac{c \phi_{est}}{2\omega}. \quad (4)$$

The same measurements are also used to estimate the amplitude as

$$\alpha_{est} = \frac{1}{2} \sqrt{(i_0 - i_\pi)^2 + \left(i_{\frac{3\pi}{2}} - i_{\frac{\pi}{2}}\right)^2}. \quad (5)$$

More detailed discussions of the basic principle of operation of time-of-flight cameras can be found in the literature [Lange and Seitz 2001; Gokturk et al. 2004; Böttgen and Seitz 2008].

3.2 Velocity via Orthogonal Heterodyne Waveforms

Direct imaging of per-pixel radial velocity information is possible with Doppler Time-of-Flight Imaging [Heide et al. 2015]. Here, the time-of-flight camera operates in *heterodyne* mode, where illumination frequency ω_g and sensor demodulation frequency ω_f are different. More precisely, the frequencies are chosen to be orthogonal within the exposure T :

$$\omega_g = o \frac{2\pi}{T} \quad \text{and} \quad \omega_f = p \frac{2\pi}{T} \quad \text{with } o, p \in \mathbb{N}, o \neq p, \quad (6)$$

For arbitrarily-long integration times, any sinusoids with different frequencies are orthogonal. However, that is not the case for finite integration times. Frequencies satisfying Eq. 6 facilitate orthogonality even for very short integration periods with closely spaced frequencies (differences of a few KHz instead of MHz). Therefore, they allow for effective usage of the AM-spectrum, which is for example important when many cameras are interfering, as well as avoiding the resolution loss associated with significantly lower frequencies.

For these particular frequency pairs, the integrated intensity measurements of static scenes are constant and do not depend on scene depth

$$i_\psi = \int_{t_0}^{t_0+T} f_\psi(t) s(t) dt = \text{const}. \quad (7)$$

However, for dynamic objects that move at a radial velocity v towards or away from the camera, the Doppler effect results in a slight frequency shift of the illumination waveform before it is demodulated on the sensor. The frequency shift is $\Delta\omega = \frac{2v}{c} \omega_g$, where c is the speed of light.

Given an intensity measurement with an orthogonal frequency pair, one computes the ratio r of this heterodyne measurement and an additional homodyne measurement as

$$\begin{aligned} r &= \frac{\int_{t_0}^{t_0+T} \cos(\omega_f t + \psi) (\alpha \cos((\omega_g + \Delta\omega)(t - \frac{2d\omega}{c})) + b) dt}{\int_{t_0}^{t_0+T} \cos(\omega_g t + \psi) (\alpha \cos((\omega_g + \Delta\omega)(t - \frac{2d\omega}{c})) + b) dt} \\ &\approx \frac{-\Delta\omega}{\omega_f - \omega_g}. \end{aligned} \quad (8)$$

This ratio image is directly proportional to the radial velocity of the imaged scene. For more detailed derivations, we refer the interested reader to [Heide et al. 2015].

3.3 Multi-source Interference for Static Scenes

A challenging problem when working with ToF camera arrays is multi-device interference (MDI). The temporally-coded illumination waveforms of each camera interfere with one another, thereby corrupting the measurements of all sensors. As an alternative, only a single light source could be employed for the array. In that case, however, parts of the imaged scene would never be illuminated, which defies the purpose of the different perspectives captured by the sensor array.

To model the interference between M light sources observed by one camera, we simply sum the contributions of each light source reflected by a Lambertian object with surface normal \mathbf{n} towards the sensor as

$$s(t) = \sum_{l=1}^M \alpha_l \cos(\omega_l t - \phi_l) (\mathbf{w}_l \cdot \mathbf{n}) + b. \quad (9)$$

Here, \mathbf{w}_l is the normalized direction of the vector pointing from the imaged scene point towards light source l . The amplitudes α_l combine the surface albedo, the brightness of the illumination, and other factors. Each light source l has its own frequency ω_l and depth phase $\phi_l = -2\omega(d_l + d)/c$, which depends on the distance between scene point and camera d and on the distance between scene point and light source d_l .

Although the amplitudes of multiple light sources additively superimpose in the sensor signal $s(t)$, they corrupt the estimated depth after demodulation (see Figure 1). Inspired by orthogonal waveform design for velocity imaging, we propose to cancel the contribution of all but one light source for each camera using mutually-orthogonal frequency pairs to mitigate MDI artifacts. For this purpose, we pick a frequency $\omega_f = p \frac{2\pi}{T}$ for sensor demodulation, as well as mutually orthogonal frequencies $l = 1 \dots M$ as $\omega_l = (p + l - 1) \frac{2\pi}{T}$ for each light source. For this particular choice of illumination frequencies, the camera operates in homodyne mode with respect to the first light source, but in (orthogonal) heterodyne mode with respect to all the other lights. The resulting

intensity measurement is

$$\begin{aligned} i_\psi &= \int_{t_0}^{t_0+T} \cos(\omega_f t + \psi) \left(\sum_{l=1}^M \alpha_l \cos(\omega_l t - \phi_l) (\mathbf{w}_l \cdot \mathbf{n}) + b \right) dt \\ &= \alpha_1 \cos(\psi - \phi_1) (\mathbf{w}_1 \cdot \mathbf{n}) + \tilde{b}. \end{aligned} \quad (10)$$

An increasing amount of light sources in this setting leads to a decreasing signal-to-ambient light ratio. Thus, this technique relies on robust ambient cancellation of the ToF sensors. In Section 5, we relax the requirement of multiple light sources having to be orthogonal and show that, in fact, any mutually different frequencies mitigate MDI artifacts.

More sophisticated multi-light-source experiments, such as illumination demultiplexing [Kadambi et al. 2014] or ToF photometric stereo [Ti et al. 2015], could be developed on top of the proposed coding techniques.

3.4 Global Illumination and Dynamic Scenes

Note that Equation 9 assumes direct-only illumination in the scene. For the more general case of direct and global illumination effects, we can model the light reaching a sensor pixel as

$$\begin{aligned} s(t) &= \underbrace{\sum_{l=1}^M \alpha_l \cos(\omega_l t - \phi_l) (\mathbf{w}_l \cdot \mathbf{n})}_{\text{direct illumination}} \\ &\quad + \underbrace{\int_{\Omega} \rho(\mathbf{w}_l, \mathbf{w}_o, t) L_l(\mathbf{w}_l, t) (\mathbf{w}_l \cdot \mathbf{n}) d\mathbf{w}_l}_{\text{global illumination}}, \end{aligned} \quad (11)$$

where ρ is a time-dependent BRDF. An interesting choice for the illumination frequencies is the slight variation of the choice above: $\omega_l = (p + l) \frac{2\pi}{T}$. In this case, the frequencies of *all* light sources are orthogonal to that of the sensor, so that all stationary objects are ignored by the imaging system. Object motion, however, can be detected, independent of whether they are directly visible, partially visible (e.g. through fog), or only indirectly visible via a diffuse reflection.

For intuitiveness, Equations 9-11 assume that the imaged scene, or at least the directly visible parts, are static throughout the exposure time. We can incorporate object motion of directly visible scene parts by extending Equation 9 and also account for multiple sensors $k = 1 \dots N$ as

$$s_k(t) = \sum_{l=1}^M \alpha_{lk} \cos \left(\omega \left(1 + \frac{\mathbf{w}_k \cdot \mathbf{v} + \mathbf{w}_l \cdot \mathbf{v}}{c} \right) \left(t - \frac{d_l + d_k}{c} \right) \right) (\mathbf{w}_l \cdot \mathbf{n}) \quad (12)$$

What is unique in this multi-sensor setup is that the 3D vector \mathbf{v} indicating both velocity and direction of motion is projected onto the line of sight of each camera \mathbf{w}_k (see Fig. 3). Theoretically, radial velocity measurements from at least three camera perspectives therefore may allow the 3D motion vector to be recovered. An additional depth map via homodyne measurements may be necessary to establish the correspondence between camera images, however. We believe that a detailed evaluation of triangulation-based approaches for time-of-flight camera arrays with coded waveforms is an interesting direction for future research. Although Equation 12 only models motion of directly visible objects, it can be generalized to global illumination by following a similar formula as Equation 11.

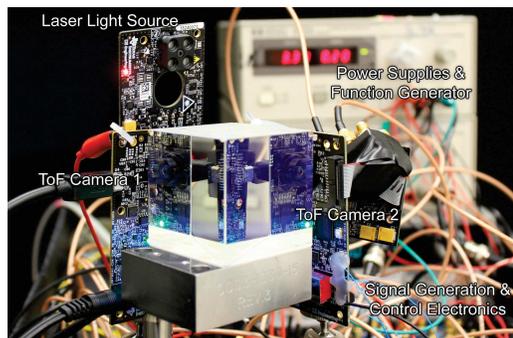


Figure 4: Prototype time-of-flight camera pair. In this configuration, two cameras are optically aligned with a beam splitter. They both share the same perspective and a single active light source, but each of the sensors is modulated at different frequencies. In this mode, we can capture homodyne and heterodyne frames simultaneously to avoid alignment artifacts of objects in motion.

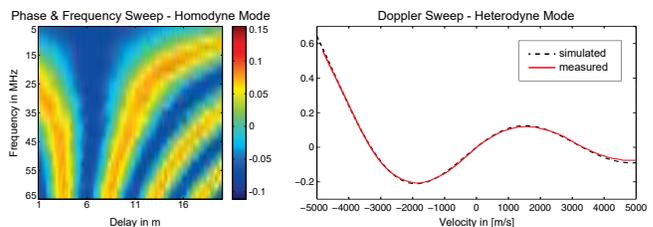


Figure 5: Measured calibration data. For the homodyne capture mode, we measure the responses for a range of different frequencies and phase delays (left); the calibration matrix looks as expected. We also validate the Doppler responses by keeping the frequency of the sensor fixed while sweeping the illumination frequency to simulate a range of different radial object velocities (right).

4 Phased ToF Camera Array Implementation

Most commercially-available time-of-flight cameras provide on-board or on-sensor signal generation. These systems are optimized for but also limited to a single task: depth imaging via phase-stepped homodyne waveforms. Precise control over the waveforms for light sources and sensor demodulation is not possible with most current-generation devices. We build on Texas Instruments' time-of-flight camera development kit (CDK OPT8241-CDK-EVM, or *Tin Tin*), and develop external control electronics for signal generation and synchronization around this CDK. An overview of the entire system is illustrated in Figure 2.

In the following, we describe our hardware system with the goal of providing the first reproducible and programmable ToF camera system. Upon publication, the list of hardware components, the capture software, and the firmware for the microcontroller will be published open source.

Time-of-Flight with Custom Waveforms *Tin Tin* is a state-of-the-art time-of-flight camera, providing a resolution of 320×240 pixels, its built-in lens offers a field of view of $74^\circ \times 59^\circ$, it is connected via USB 2.0 to a host computer, and it provides easily accessible connections for external signal control and synchronization. The internal signal generation can be disabled via appropriate register settings and the raw sensor data is read out from the host computer via the provided USB interface. *Tin Tin* has a built-in laser light engine. At 850 nm, the diffused laser diodes offer a

pulse repetition frequency of 12-80 MHz for square waves at 50% duty cycle. Both sensor and light source can be controlled with digital waveforms from external sources. The on-board light engine can also be disabled and replaced by a custom light source, such as a projector which could provide spatio-temporally structured illumination [O’Toole et al. 2014].

The waveform-generation circuitry in our system is primarily built around Direct Digital Synthesis (DDS). Compared to FPGA-based signal generation (e.g., [Kadambi et al. 2013]) DDS has demonstrated more flexibility and more precise waveform control in previous time-of-flight systems [Heide et al. 2013; Heide et al. 2015]. In particular, heterodyne frequency pairs have extreme requirements on frequency precision that we have so far only been able to reliably generate via DDS. We use an Analog Devices AD9959 four-channel DDS to generate sinusoidal waveforms with independent control over frequency, phase, and amplitude. Using these four channels, we can control either two ToF cameras with their respective sensors and light sources independently, three camera sensors that share a single light source, or a single sensor with three light sources. Each sensor or light source requires their own channel. The DDS generates analog sinusoidal waveforms using a built-in digital-to-analog converter (DAC). The modulation signals for the camera, however, need to be digital for both illumination and sensor demodulation. To digitize the analog DDS waveform, a clock fan-out buffer (Analog Devices AD9513) is used as a high-speed comparator.

We use an ARM Cortex M4 STM32F407VGT6 microcontroller to synchronize the exposures of all cameras and control the waveforms on the DDS. The microcontroller acts as the master for all connected devices. Each time-of-flight camera is set to slave mode and then waits for the microcontroller to trigger the beginning and end of the exposure. Although the waveforms are generated by the DDS and digitized by the comparator, the microcontroller also controls phase and frequency settings of all four DDS channels and updates them between all captured frames. The cameras actually distinguish between frames and *quads*. Each frame comprises four quads, which are normally used for the four phase-stepped homodyne frames. Our microcontroller sets the frequencies and phases for all four quads on the DDS at the beginning of each frame. The frequency and phase settings of each quad in every channel can be controlled individually. Phase stepping is not restricted to the conventional $\frac{\pi}{2}$ steps. The raw readout mode of the camera SDK reads out all four quads at the same time through the USB interface.

An additional high speed buffer chip (TI SNB8LVCF00A) acts as an exposure gating mechanism between the output of the comparator and the sensor. This gating mechanism disconnects the camera end completely from the signal generation circuit to ensure that no residual signals reach the sensor. Without this gating mechanism, image artifacts are observed.

Figures 1 and 4 show different configurations of our setup. We can either run the cameras in light field mode, where each sensor observes the scene from different perspectives (Fig. 1) or we can optically align them using a beam splitter so that they share the same perspective (Fig. 4). Each configuration is useful for different applications, as discussed in Section 5.

Camera Calibration To verify correct operation of the externally-controlled camera system, we simulate and measure the intensity responses for both homodyne and heterodyne operation (see Fig. 5). The matrix in the left column shows a sweep over modulation frequencies and delays for the homodyne setting. This matrix matches our expectations and is comparable to data reported in previous work (e.g., [Heide et al. 2013]). Although the light source is tested for up to 80 MHz and the sensor to about 95 MHz,

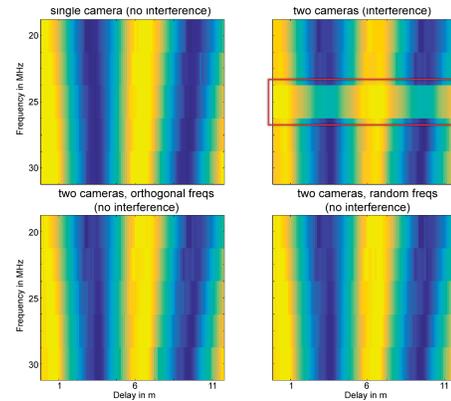


Figure 6: Multi-device interference. The top left shows the reference phase and frequency sweep response of a single camera. Two ToF cameras, each running in homodyne mode, interfere at one particular frequency with one another (top right). Choosing an orthogonal frequency pair for any setting eliminates multi-device interference (bottom left), but the same effect is also achieved by simply choosing two different frequencies (bottom right).

we can currently only drive waveforms reliably up to 65 MHz. We expect that this limitation can be lifted with a better board design, which should allow us to reach the design frequency of the sensor in the future. We also calibrate the intensity responses for the heterodyne mode. In this case, we pick a fixed frequency of 45,000 cycles per exposure time ($T = 1,970 \mu s$) for the sensor and vary the frequency of the light source. This simulates objects recorded with different radial velocities. Again, we observe a close match between predicted and measured data.

Multi-camera Synchronization and Interference For multi-camera synchronization, we set all connected sensors to slave mode, and control the exposures from the microcontroller. Running the current hardware system with multiple cameras is relatively straightforward: the same microcontroller controls all camera exposures and the DDS boards, which supports up to three sensors sharing the same light source. The system is easily extended to more than four channels by using additional DDS boards, comparators, and gating chips. All of these could be controlled by the same microcontroller. However, all DDS boards would have to share the same clock.

In Figure 6, we evaluate the interference patterns between multiple time-of-flight cameras. On the top left, we show the phase and frequency sweep from 20 to 30 MHz for a single camera. This serves as the reference plot. On the top right, we activate a second camera with its own light source running at a fixed frequency of 25 MHz. We observe that no interference between the cameras occurs unless both use the same frequencies. As predicted, choosing the orthogonal frequency for any of the swept frequencies eliminates interference (bottom left). However, we can also eliminate interference by simply choosing *any* mutually different frequency pairs. In the example shown in the bottom right, we set the fixed frequency of one camera to 18 MHz and sweep the other; no interference is visible. Therefore, multi-device interference in time-of-flight camera arrays can simply be overcome by choosing different but not necessarily orthogonal frequencies for all connected devices.

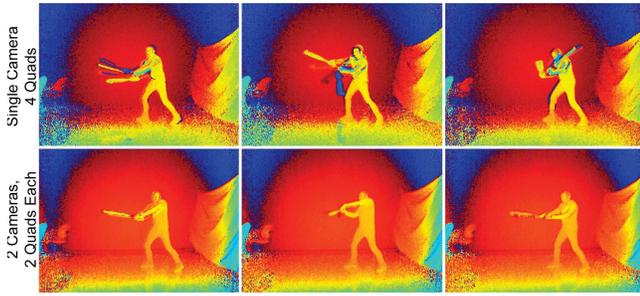


Figure 7: Conventional depth imaging with a single camera (top row) and with a phased camera pair (bottom row). Estimating the depth map requires four phase-stepped measurements (quads). With any single camera (top row), these are recorded in sequence, which introduces motion artifacts. With multiple phased time-of-flight cameras, these measurements can be recorded in parallel. We demonstrate this for a phased camera pair (bottom row), where each of the sensors captures two of the four quads. Motion artifacts are significantly reduced, but only with four optically-aligned and phased ToF cameras can these artifacts be fully corrected.

5 Applications

In this section, we outline a variety of applications for multi-camera time-of-flight systems. First, we demonstrate that our system is capable of capturing range and radial velocity data. We also demonstrate new applications, including multi-device interference cancellation, fast range and velocity estimation without motion artifacts as well as non-line-of-sight motion detection and de-scattering of moving scenes.

Fast Range Imaging with Phased ToF Cameras As outlined by Equation 4, four phase-stepped measurements are necessary to estimate a depth map with a time-of-flight camera. The most intuitive application for a phased time-of-flight camera arrays would be to capture all of these measurements in parallel. This is beneficial for mitigating motion artifacts observed for dynamic scenes. For example, we show a person swinging a baseball bat in Figure 7. Our ToF cameras record 60 frames per second (fps), and they periodically cycle through the four different phase settings. In practice, this results in a frame rate of 15 fps, but the individual sub-frames or *quads* are not aligned for fast-moving objects (Fig. 7, top row). By capturing two pairs of different phase settings with two phased cameras effectively doubles the camera framerate and mitigates motion artifacts (Fig. 7, bottom row).

Range and Velocity Imaging via D-ToF Imaging As discussed in Section 3.2, radial velocity imaging requires the ratio between a heterodyne image captured with orthogonal frequencies and a homodyne image to be computed. In previous work [Heide et al. 2015], this was achieved with a single time-of-flight camera by recording the two measurements in an alternating manner, thereby introducing alignment artifacts of quickly moving objects. We verify this capture mode with our system in Figure 8 and make two interesting observations. First, whenever we measure radial velocity in the scene, there is usually also lateral motion. This leads to image alignment artifacts due to the sequential nature of the recording process and is expected but cannot be avoided. Second, we observe significantly better image quality than what was reported in previous work [Heide et al. 2015]. Not only is the image resolution four times higher, but the signal-to-noise-ratio of the heterodyne frames is much higher. As opposed to previous work, we do not have to apply any denoising to the raw frames. Although that strategy could

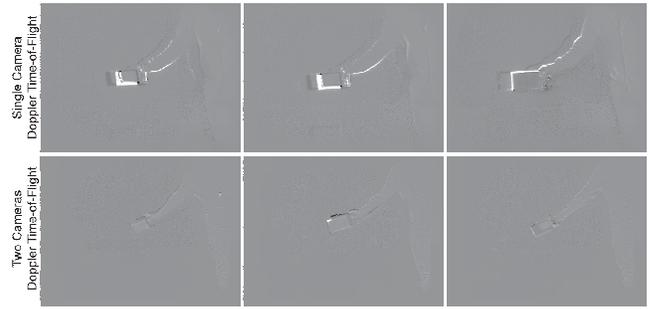


Figure 8: Radial velocity frames from Doppler imaging for a single time-of-flight camera similar to [Heide et al. 2015] (top) and our two camera solution (bottom). Temporal alignment artifacts between the homodyne and the heterodyne frame for the single camera are not digitally corrected and show up as artifacts around depth discontinuities. By using two cameras, each capturing either homodyne or heterodyne mode only, we can mitigate these artifacts. Residual frame differences are from a slight amount of parallax between the two cameras.

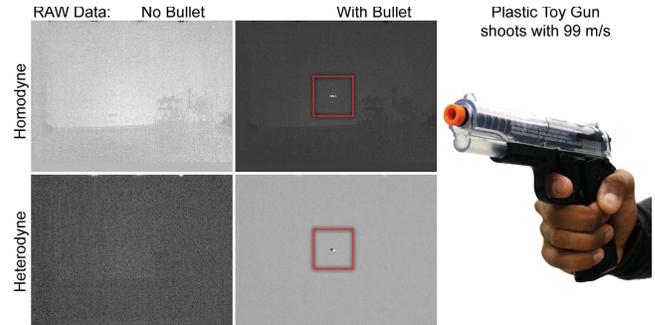


Figure 9: Example scene of a flying bullet. Using the beam splitter setup shown in Figure 4, we simultaneously capture homodyne and heterodyne frames of a static scene (left) and of the same scene with a bullet flying at about 99 m/s (center). The heterodyne frame of the moving projectile only shows the intensity response of the Doppler frequency shift. The two images can be combined according to Eq. 8 to mitigate temporal alignment issues of Doppler velocity captured with a single sensor. Note that the homodyne and heterodyne frames are normalized and not on the same intensity scale.

further improve the quality of the reconstructions, modern denoising techniques, such as non-local means [Buares et al. 2005], impose a significant computational burden.

A phased pair of time-of-flight cameras can partially overcome the alignment artifacts observed in the sequential capture mode. In this case, the two cameras share a single light source. One of the sensors uses the same modulation frequency as the light source (homodyne mode) and the other sensor uses an orthogonal frequency (heterodyne mode). The exposure times of each camera are synchronized. Several different configurations for such a camera pair exist. When the objects of interest are far away, the cameras could be placed next to each other (see Fig. 1) or they could otherwise be optically combined with a beam splitter (see Fig. 4). In either case, possible misalignment between the two streams can be reasonably well corrected with a global homography. With an increasing baseline between the cameras, the misalignment would be depth-dependent and could not be easily corrected.

We demonstrate results for both optical configurations in Figures 8

and 9, respectively. Figure 8 compares similar types of motion captured with the sequential mode discussed above and also with a phased camera pair. Each of the cameras continuously streams raw frames with either homodyne-only or heterodyne-only mode. The alignment artifacts observed in the sequential mode are mitigated by the camera pair, because this approach to Doppler Time-of-Flight Imaging is closer to being instantaneous³. Small amounts of misalignment remain and are due to parallax between the cameras, which can only be perfectly corrected for a single depth with a homography.

Figure 9 shows a captured bullet in flight captured by the beam splitter setup. This projectile is fired from a plastic toy gun with about 99 m/s, as listed by the manufacturer. When only the static background is visible (left column), the intensity response of the homodyne frame is related to the scene depth and the orthogonal heterodyne frame observes no response. When the bullet is captured within an exposure, the homodyne frame contains distorted depth information and the heterodyne frame exhibits a strong intensity response at the correct location.

Multi-Camera Interference Cancellation Interference cancellation for multiple depth cameras has been a topic of interest in the context of structured illumination [Butler et al. 2012; Maimone and Fuchs 2012] and also for time-of-flight cameras [Castaneda et al. 2014; Li et al. 2015]. As discussed in prior work and experimentally verified in Section 4, multiple time-of-flight cameras running at the same frequency will interfere with one another. As shown in Figure 1 and in the supplemental video, multi-device interference leads to a low-frequency beating pattern that distorts the recovered depth map for each of the connected cameras. Whereas Castaneda et al. proposed to capture all possible combinations of activated light sources, Lianhua et al. proposed to capture and average more than 100 frames to statistically mitigate multi-device interference. Neither option seems feasible for dynamic scenes. As demonstrated in Figure 6, however, there is a simple solution to this problem: use different (homodyne) frequencies for each camera / light source pair. The waveforms do not necessarily have to be orthogonal for this technique. Figure 1 demonstrates that this simple solution works robustly for a test case with two simultaneously-operating time-of-flight cameras, each running at a different frequency in homodyne mode.

Motion Detection for Non-Line-of-Sight Non-line-of-sight motion detection is a topic that has recently gained a lot of interest [Garipey et al. 2016]. Instead of having to rely on expensive and delicate single photon avalanche diodes, as Garipey et al. do, we utilize inexpensive time-of-flight cameras with orthogonal frequencies. The idea motivating this application was discussed in Section 3.4: when using orthogonal frequencies for the illumination and sensor demodulation waveforms, all light that is reflected from static objects towards the camera is optically gated out and appears as an ambient component that the ToF camera automatically removes. We already demonstrated that this mechanism facilitates radial velocity imaging of directly-visible objects, but the same technique conceptually also allows for motion in the global illumination components of a scene to be separated from direct reflections of static objects (see Sec. 3.4). Compared to non-line-of-sight shape reconstruction [Velten et al. 2012; Heide et al. 2014a], motion detection is much simpler and does not require a costly inverse problem to be solved. Compared to Doppler *velocity measurement* ([Heide et al. 2015], Eq. 8), *motion detection* just requires a single ToF camera and one or more heterodyne light sources. We simply look for intensity changes in the heterodyne images that indicate

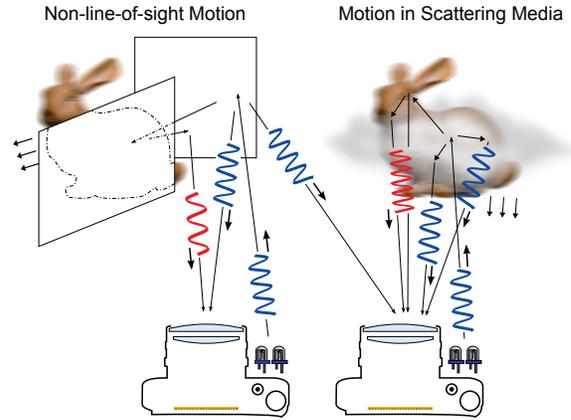


Figure 10: When using orthogonal frequencies in time-of-flight imaging, light that is directly or indirectly reflected from static objects is gated out by the sensor and not visible in the captured images. Objects in motion distort the temporal waveforms, which is observed as intensity changes. This effect applies to directly visible objects, but also to objects that are outside the direct line-of-sight or objects veiled by scattering.

the presence of motion in indirectly reflected illumination but we do not aim to recover their shape. While a single light source is sufficient, multiple lights with orthogonal frequencies can be used to boost the sensitivity to a larger range of object velocities. This mode of operation is easily supported by our hardware system. The concept of motion detection not only applies to motion of objects outside the direct line-of-sight but also of objects that are veiled by scattering. We illustrate this concept in Figure 10.

We evaluate non-line-of-sight motion detection in Figure 11. Here, the camera is pointed towards a big, flat, white, diffuse target (outside the photograph, on the left). We select orthogonal frequencies for the light source and the sensor of this single-camera setup and rapidly move a diffuse or a specular object behind the camera. This is a three-bounce global illumination experiment. The diffuse object reflects light back to the directly-visible target in all directions, thereby causing slight, yet visible, low-frequency changes in the heterodyne image. When we move a specular object, more indirectly-reflected light is focused on certain parts of the big target than on others. The effect of motion for the specular non-line-of-sight object is well visible.

De-scattering of Dynamic Scenes Finally, we also demonstrate the frequency gating mechanism described in the previous section for de-scattering of dynamic scenes. As opposed to recent work on computational de-scattering [Heide et al. 2014b], we do not aim to recover the shape of the veiled 3D object computationally. Our approach is more similar to the recently-proposed depth-selective coding technique proposed by Tadano et al. [2015] in that we engineer waveforms to achieve a certain goal directly using through waveform correlation on the sensor. Tadano et al. proposed to use waveforms that would optically gate out certain path lengths. We build on the idea of orthogonal frequencies with the motivation that a scattering medium is mostly static and therefore have no correlation with the camera modulation signal. Modern time-of-flight cameras are very effective at suppressing uncorrelated light components, effectively gating out the directly reflected light. Only those light rays that directly interacted with a dynamic object inside the participating medium and that are then scattered back to the sensor will be observed. We note, however, that even the suppressed component

³Instantaneity in this context is defined on a per-frame basis.

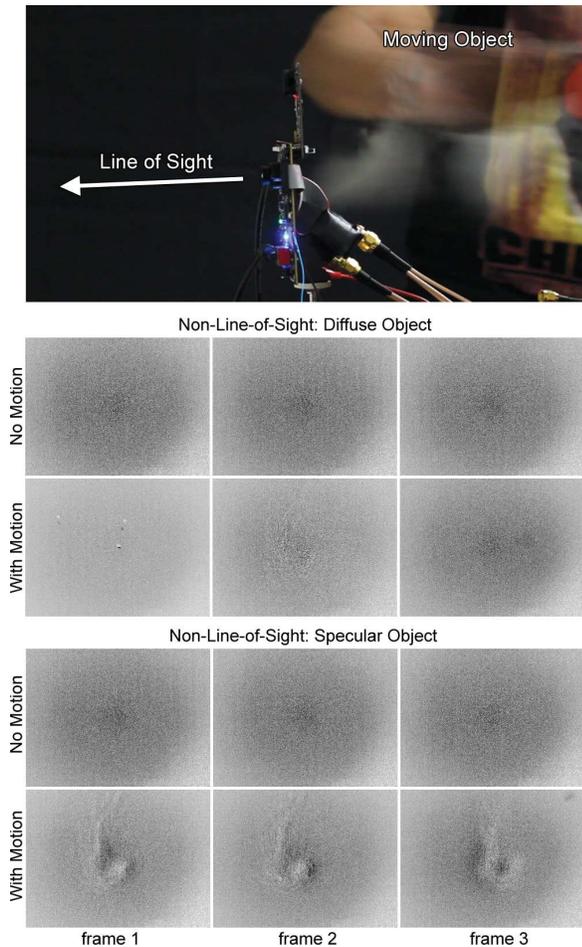


Figure 11: *Non-line-of-sight (NLOS) motion detection with a single ToF camera and a heterodyne light source. The camera and light source are directly pointed towards a flat, diffuse target (top, target not shown) whereas an object moves at high speed behind the camera. This setup demonstrates three-bounces of global illumination. We experiment with a diffuse NLOS object (center) and a specular object (bottom). Neither object is visible in any amplitude or homodyne images. However, these heterodyne images clearly demonstrate that Doppler frequency shift of the global illumination changes with non-line-of-sight motion. This effect is more subtle for a diffuse NLOS object than for a specular one, but it is visible in low-frequency intensity changes of the heterodyne images.*

contributes to the photon shot noise, so that large amounts of scattering show up as increased levels of Skellam noise in the measured motion images.

We demonstrate this effect in Figure 12, where a fog machine is used to partially veil the person of interest. In both cases, the person is rapidly moving an object with their hands but due to scattering, this object is not visible to a conventional camera or, as shown here, in the amplitude image of the time-of-flight camera. The gating mechanism of the proposed waveforms removes a significant amount of scattered light, such that the veiled dynamic parts of the scene become visible in the heterodyne images without any additional processing. We envision this technique to be able to assist autonomous and remotely controlled vehicles in poor weather conditions, under water, or in challenging environments.

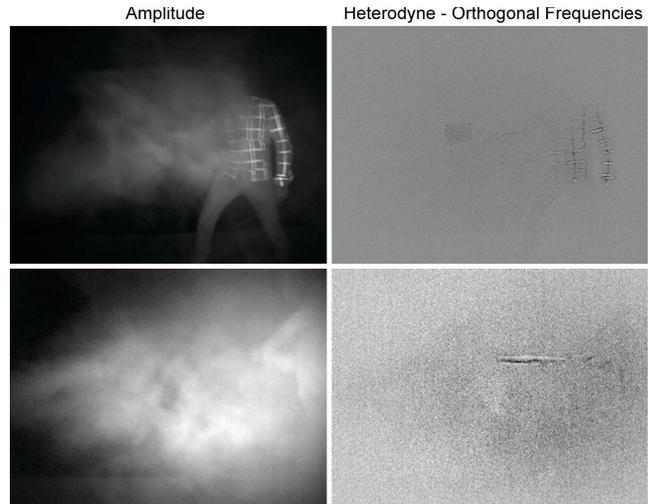


Figure 12: *Seeing motion through scattering media. We use a fog machine to obscure the direct line-of-sight from the camera to the moving object. In the amplitude images, the moving parts of the scene are completely veiled by scattering (left). The frequency gating mechanism of the orthogonal frequencies used for illumination and sensor waveforms separates all direct and global illumination that was not affected by motion from that affected by motion. This leads to a de-scattering effect of the dynamic scene parts (right).*

6 Discussion

In summary, we develop a phased multi-camera time-of-flight system that serves as a platform for evaluating a number of computational imaging applications. We plan on making this hardware system, the required firmware, as well as software for calibration and data readout freely-available online. We propose custom waveforms for novel time-of-flight applications, such as multi-device interference cancellation, de-scattering dynamic scenes, and non-line-of-sight motion detection via frequency gating. We also show how to improve several existing applications, including fast range imaging as well as artifact-free radial velocity imaging.

Extension to More Sensors and Light Sources Although currently limited to four channels, the proposed system design is extensible to support more sensors and light sources. Extending the system primarily requires additional modulation channels that are independently configurable by the MCU and a mechanism to read data from all sensors in the array. Modulation channel capacity can be increased by adding multiple 4-channel DDS (AD9959) integrated circuits (ICs) to the system. The DDS ICs need to be clocked by the same reference clock and this can be accomplished using a clock distribution IC (AD9510) with a high frequency (≈ 500 MHz) SAW oscillator. Proper layout considerations like matched trace lengths need to be taken into account to prevent clock skew. The DDS ICs would also have to be synchronized by either daisy chaining SYNC_OUT/SYNC_IN lines or by using the SYNC_OUT from the first IC to provide SYNC_IN to the rest using a clock distribution IC similar to the one mentioned above.

The microcontroller used in our system consists of two separate SPI peripherals which will allow for completely independent communication with up to two DDS ICs. To extend beyond that (8+ channels), a multi-slave SPI configuration can be adopted where the Slave Select (SS) line signals the IC that needs to be configured on the shared bus. Theoretically, there is no fundamental limit to

the number of modulation channels that the proposed design could support. However, adding more channels, and therefore DDS ICs, does increase communication overhead in the order of $\approx 10\mu s$ per DDS as they need to be configured separately using SPI commands at the beginning of exposure.

With channel capacity extended, readout from multiple sensors in the array could be achieved using separate USB ports with independent bus controllers on the host PC. Upon being bandwidth limited, additional Host PCs could be used for readout. The readout process already uses a frame identifier which could also be used to solve correspondence issues with data acquired over a distributed memory readout system.

Future Work In addition to extending the system to support more sensors and light sources, we plan on moving the benchtop system to a custom printed circuit board. This would make our electronics platform portable and most likely also overcome the current frequency limit of 65 MHz. We would like to extend the system to more cameras and light sources and validate all demonstrated applications in outdoor settings. Further, we would like to experiment with high dynamic range capture modes where each camera in the array records different exposures and we would like to evaluate our hardware system for recently-proposed applications in non-line-of-sight imaging, explore new directions in direct/global illumination decomposition, and “upgrade” the light sources to programmable spatio-temporally-coded projectors.

7 Conclusion

Computational time-of-flight imaging is an emerging, yet active research area. We explore new directions in multi-device interference cancellation, optimized range and velocity imaging, and direct/global illumination separation for dynamic scenes. With this work, we lay the foundations of computational imaging with phased time-of-flight camera arrays. We hope that our reproducible hardware platform along with the diversity of applications we explore stimulates further research on multi-sensor, multi-light source coded computational photography.

Acknowledgements

This work was generously supported by Texas Instruments, the National Science Foundation under grant IIS 1553333, by the NSF/Intel Partnership on Visual and Experiential Computing (NSF IIS 1539120), by the KAUST Office of Sponsored Research through the Visual Computing Center CCF grant, and by Intuitive Surgical.

References

BAMJI, C., O’CONNOR, P., ELKHATIB, T., MEHTA, S., THOMPSON, B., PRATHER, L., SNOW, D., AKKAYA, O., DANIEL, A., PAYNE, A., AND M. FENTON, T. P., AND CHAN, V. 2015. A 0.13 μm CMOS System-on-Chip for a 512 x 424 Time-of-Flight Image Sensor With Multi-Frequency Photo-Demodulation up to 130 MHz and 2 GS/s ADC. *IEEE Journal of Solid-State Circuits* 50, 1, 303–319.

BHANDARI, A., KADAMBI, A., WHYTE, R., BARSII, C., FEIGIN, M., DORRINGTON, A., AND RASKAR, R. 2014. Resolving multipath interference in time-of-flight imaging via modulation frequency diversity and sparse regularization. *Optics Letters* 39, 1705–1708.

BUADES, A., COLL, B., AND MOREL, J.-M. 2005. A non-local algorithm for image denoising. In *Proc. IEEE CVPR*, vol. 2.

BUEHLER, C., BOSSE, M., McMILLAN, L., GORTLER, S., AND COHEN, M. 2001. Unstructured lumigraph rendering. In *Proc. SIGGRAPH*, 425–432.

BUTLER, D., IZADI, S., HILLIGES, O., MOLYNEAUX, D., HODGES, S., AND KIM, D. 2012. Shake’n’sense: reducing interference for overlapping structured light depth cameras. In *Proc. ACM UIST*, 1933–1936.

BÜTTGEN, B., AND SEITZ, P. 2008. Robust optical time-of-flight range imaging based on smart pixel structures. *IEEE Trans. Circuits and Systems* 55, 6, 1512–1525.

CARRANZA, J., THEOBALT, C., MAGNOR, M. A., AND SEIDEL, H.-P. 2003. Free-viewpoint video of human actors. *ACM Trans. Graph. (SIGGRAPH)* 22, 3, 569–577.

CASTANEDA, V., MATEUS, D., AND NAVAB, N. 2014. Stereo time-of-flight with constructive interference. *IEEE Trans. PAMI* 36, 7, 1402–1413.

DEBEVEC, P., HAWKINS, T., TCHOU, C., DUIKER, H.-P., SAROKIN, W., AND SAGAR, M. 2000. Acquiring the reflectance field of a human face. In *Proc. SIGGRAPH*, 145–156.

DONG, S., HORSTMAYER, R., SHIRADKAR, R., GUO, K., OU, X., BIAN, Z., XIN, H., AND ZHENG, G. 2014. Aperture-scanning fourier ptychography for 3d refocusing and super-resolution macroscopic imaging. *Optics Express* 22, 11, 13586–99.

DORRINGTON, A., GODBAZ, J., CREE, M., PAYNE, A., AND STREETER, L. 2011. Separating true range measurements from multi-path and scattering interference in commercial range cameras. In *Proc. Electronic Imaging*.

FREEDMAN, D., KRUPKA, E., SMOLIN, Y., LEICHTER, I., AND SCHMIDT, M. 2014. Sra: fast removal of general multipath for tof sensors. In *Proc. ECCV*.

FUCHS, S. 2010. Multipath interference compensation in time-of-flight camera images. In *Proc. ICPR*.

GALL, J., HO, H., IZADI, S., KOHLI, P., REN, X., AND YANG, R. 2014. Towards solving real-world vision problems with rgb-d cameras. In *CVPR Tutorial*.

GARIEPY, G., TONOLINI, F., AND JONATHAN LEACH, R. H., AND FACCIO, D. 2016. Detection and tracking of moving objects hidden from view. *Nature Photonics Letters* 10, 23–26.

GOKTURK, S., YALCIN, H., AND BAMJI, C. 2004. A time-of-flight depth sensor - system description, issues and solutions. In *Proc. CVPR*, 35–35.

GORTLER, S. J., GRZESZCZUK, R., SZELISKI, R., AND COHEN, M. F. 1996. The lumigraph. In *Proc. SIGGRAPH*.

HANSARD, M., LEE, S., CHOI, O., AND HORAUD, R. 2012. *Time of Flight Cameras: Principles, Methods, and Applications*. Springer.

HEIDE, F., HULLIN, M. B., GREGSON, J., AND HEIDRICH, W. 2013. Low-budget transient imaging using photonic mixer devices. *ACM Trans. Graph. (SIGGRAPH)* 32, 4, 45:1–45:10.

HEIDE, F., XIAO, L., HEIDRICH, W., AND HULLIN, M. B. 2014. Diffuse mirrors: 3D reconstruction from diffuse indirect illumination using inexpensive time-of-flight sensors. In *Proc. CVPR*.

- HEIDE, F., XIAO, L., KOLB, A., HULLIN, M. B., AND HEIDRICH, W. 2014. Imaging in scattering media using correlation image sensors and sparse convolutional coding. *Optics Express* 22, 21, 26338–26350.
- HEIDE, F., HEIDRICH, W., HULLIN, M., AND WETZSTEIN, G. 2015. Doppler Time-of-Flight Imaging. *ACM Trans. Graph. (SIGGRAPH)*, 4.
- HOLLOWAY, J., SALMAN ASIF, M., SHARMA, M. K., MATSUDA, N., HORSTMAYER, R., COSSAIRT, O., AND VEERARAGHAVAN, A. 2015. Toward Long Distance, Sub-diffraction Imaging Using Coherent Camera Arrays. *ArXiv 1510.08470*.
- JAYASURIYA, S., PEDIREDLA, A., SIVARAMAKRISHNAN, S., MOLNAR, A., AND VEERARAGHAVAN, A. 2015. Depth fields: Extending light field techniques to time-of-flight imaging. In *Proc. 3DV*, 1–9.
- JIMENEZ, D., PIZARRO, D., MAZO, M., AND PALAZUELOS, S. 2012. Modelling and correction of multipath interference in time of flight cameras. In *Proc. CVPR*.
- KADAMBI, A., WHYTE, R., BHANDARI, A., STREETER, L., BARSİ, C., DORRINGTON, A., AND RASKAR, R. 2013. Coded time of flight cameras: sparse deconvolution to address multipath interference and recover time profiles. *ACM Trans. Graph. (SIGGRAPH Asia)* 32, 6.
- KADAMBI, A., BHANDARI, A., WHYTE, R., DORRINGTON, A., AND RASKAR, R. 2014. Demultiplexing Illumination via Low Cost Sensing and Nanosecond Coding. In *Proc. ICCP*.
- KIM, S. K., KANG, B., HEO, J., JUNG, S.-W., AND CHOI, O. 2014. Photometric stereo-based single time-of-flight camera. *Optics Letters* 39, 1, 166–169.
- KIRMANI, A., HUTCHISON, T., DAVIS, J., AND RASKAR, R. 2009. Looking around the corner using transient imaging. In *Proc. ICCV*, 159–166.
- LANGE, R., AND SEITZ, P. 2001. Solid-state time-of-flight range camera. *IEEE J. Quantum Electronics* 37, 3, 390–397.
- LEVOY, M., AND HANRAHAN, P. 1996. Light field rendering. In *Proc. SIGGRAPH*, 31–42.
- LI, L., XIANG, S., YANG, Y., AND YU, L. 2015. Multi-camera interference cancellation of time-of-flight (tof) cameras. In *Proc. IEEE ICIP*, 556–560.
- MAIMONE, A., AND FUCHS, H. 2012. Reducing interference between multiple structured light depth sensors using motion. In *Proc. VR*.
- MATUSIK, W., BUEHLER, C., RASKAR, R., GORTLER, S. J., AND MCMILLAN, L. 2000. Image-based visual hulls. In *Proc. SIGGRAPH*, 369–374.
- MCCANDLESS, S. W., AND JACKSON, C. R. 2004. Principles of synthetic aperture radar. In *AR Marine Users Manual*, J. Fagerberg, D. C. Mowery, and R. R. Nelson, Eds. NOAA, ch. 1, 11.
- NAIK, N., ZHAO, S., VELTEN, A., RASKAR, R., AND BALA, K. 2011. Single view reflectance capture using multiplexed scattering and time-of-flight imaging. *ACM Trans. Graph. (SIGGRAPH Asia)* 30, 6, 171:1–171:10.
- NAIK, N., KADAMBI, A., RHEMANN, C., IZADI, S., RASKAR, R., AND KANG, S. 2015. A light transport model for mitigating multipath interference in tof sensors. In *Proc. CVPR*.
- O'TOOLE, M., HEIDE, F., XIAO, L., HULLIN, M. B., HEIDRICH, W., AND KUTULAKOS, K. N. 2014. Temporal frequency probing for 5d transient analysis of global light transport. *ACM Trans. Graph. (SIGGRAPH)* 33, 4, 87:1–87:11.
- PAYNE, A., JONGENELEN, A., DORRINGTON, A., CREE, M., AND CARNEGIE, D. 2009. Multiple Frequency Range Imaging to Remove Measurement Ambiguity. In *Proc. Optical 3-D measurement techniques IX*.
- PETERS, C., KLEIN, J., HULLIN, M. B., AND KLEIN, R. 2015. Solving trigonometric moment problems for fast transient imaging. *ACM Trans. Graph. (SIGGRAPH Asia)* 34, 6.
- RANDER, P., NARAYANAN, P. J., AND KANADE, T. 1997. Virtualized reality: Constructing time-varying virtual worlds from real events. In *Proc. IEEE Visualization*, 277–283.
- SHOTTON, J., FITZGIBBON, A., COOK, M., SHARP, T., FINOCCHIO, M., MOORE, R., KIPMAN, A., AND BLAKE, A. 2011. Real-time human pose recognition in parts from single depth images. In *Proc. CVPR*.
- TADANO, R., PEDIREDLA, A. K., AND VEERARAGHAVAN, A. 2015. Depth selective camera: A direct, on-chip, programmable technique for depth selectivity in photography. In *Proc. IEEE ICCV*.
- TI, C., YANG, R., DAVIS, J., AND PAN, Z. 2015. Simultaneous Time-of-Flight Sensing and Photometric Stereo With a Single ToF Sensor. In *Proc. CVPR*.
- VELTEN, A., WILLWACHER, T., GUPTA, O., VEERARAGHAVAN, A., BAWENDI, M., AND RASKAR, R. 2012. Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging. *Nat Commun* 745, 3.
- VELTEN, A., WU, D., JARABO, A., MASIA, B., BARSİ, C., JOSHI, C., LAWSON, E., BAWENDI, M., GUTIERREZ, D., AND RASKAR, R. 2013. Femto-photography: Capturing and visualizing the propagation of light. *ACM Trans. Graph. (SIGGRAPH)* 32, 4, 44:1–44:8.
- WILBURN, B., JOSHI, N., VAISH, V., TALVALA, E.-V., ANTUNEZ, E., BARTH, A., ADAMS, A., HOROWITZ, M., AND LEVOY, M. 2005. High performance imaging using large camera arrays. *ACM Trans. Graph. (SIGGRAPH)* 24, 3, 765–776.
- WOODHAM, R. J. 1980. Photometric method for determining surface orientation from multiple images. *Optical Engineering* 19, 1.
- WU, D., WETZSTEIN, G., BARSİ, C., WILLWACHER, T., O'TOOLE, M., NAIK, N., DAI, Q., KUTULAKOS, K., AND RASKAR, R. 2012. Frequency analysis of transient light transport with applications in bare sensor imaging. In *Proc. ECCV*.