Doppler Time-of-Flight Imaging

Felix Heide[1,2,3] Wolfgang Heidrich[2,1] Matthias Hullin[4] Gordon Wetzstein[3]
[1]University of British Columbia [2]KAUST [3]Stanford University [4]University of Bonn

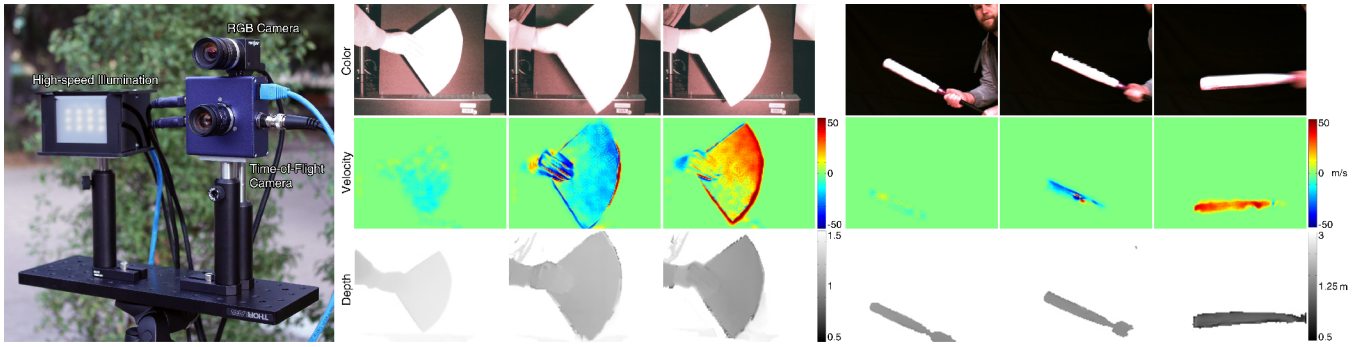

Figure 1: *We introduce a new computational imaging system that allows for metric radial velocity information to be captured instantaneously for each pixel (center row). For this purpose, we design the temporal illumination and modulation frequencies of a time-of-flight camera (left) to be orthogonal within its exposure time. The Doppler effect of objects in motion is then detected as a frequency shift of the illumination, which results in a mapping from object velocity to recorded pixel intensity. By capturing a few coded time-of-flight measurements and adding a conventional RGB camera to the setup, we demonstrate that color, velocity, and depth information of a scene can be recorded simultaneously. The results above show several frames of two video sequences. For each example, the left-most frame shows a static object (velocity map is constant), which is then moved towards (positive radial velocity) or away from (negative velocity) the camera.*

Abstract

Over the last few years, depth cameras have become increasingly popular for a range of applications, including human-computer interaction and gaming, augmented reality, machine vision, and medical imaging. Many of the commercially-available devices use the time-of-flight principle, where active illumination is temporally coded and analyzed in the camera to estimate a per-pixel depth map of the scene. In this paper, we propose a fundamentally new imaging modality for all time-of-flight (ToF) cameras: per-pixel radial velocity measurement. The proposed technique exploits the Doppler effect of objects in motion, which shifts the temporal illumination frequency before it reaches the camera. Using carefully coded illumination and modulation frequencies of the ToF camera, object velocities directly map to measured pixel intensities. We show that a slight modification of our imaging system allows for color, depth, and velocity information to be captured simultaneously. Combining the optical flow computed on the RGB frames with the measured metric radial velocity allows us to further estimate the full 3D metric velocity field of the scene. The proposed technique has applications in many computer graphics and vision problems, for example motion tracking, segmentation, recognition, and motion deblurring.

CR Categories: I.3.3 [Computer Graphics]: Picture/Image Generation—Digitizing and scanning

Keywords: computational photography, time-of-flight

1 Introduction

Pioneers of photography, including Eadweard Muybridge and Harold "Doc" Edgerton, advanced imaging technology to reveal otherwise invisible motions of high-speed events. Today, understanding the motion of objects in complex scenes is at the core of computer vision, with a wide range of applications in object tracking, segmentation, recognition, motion deblurring, navigation of autonomous vehicles, and defense. Usually, object motion or motion parallax are estimated via optical flow [Horn and Schunck 1981]: recognizable features are tracked across multiple video frames. The computed flow field provides the basis for many computer vision algorithms, including depth estimation. Unfortunately, optical flow is computationally expensive, fails for untextured scenes that do not contain good features to track, and it only measures 2D lateral motion perpendicular to the camera's line of sight. Further, the unit of optical flow is pixels; metric velocities cannot be estimated unless depth information of the scene is also available. For the particular application of depth estimation, many limitations of optical flow estimation can be overcome using active illumination, as done by most structured illumination and time-of-flight (ToF) cameras. With the emergence of RGB-D imaging, for example facilitated by Microsoft's Kinect One[1], complex and untextured 3D scenes can be tracked by analyzing both color and depth information, resulting in richer visual data that has proven useful for many applications.

In this paper, we introduce a new approach to directly imaging radial object velocity. Our approach analyzes the Doppler effect in time-of-flight cameras: object motion towards or away from the camera shifts the temporal illumination frequency before it is recorded. Conventional time-of-flight cameras encode phase information (and therefore scene depth) into intensity measurements. Instead, we propose Doppler Time-of-Flight (D-ToF) as a new imaging mode, whereby the change of illumination frequency (corresponding to radial object velocity) is directly encoded into the measured intensity. The required camera hardware is the same as for

[1] microsoft.com/en-us/kinectforwindows/

conventional time-of-flight imaging, but illumination and modulation frequencies are carefully designed. We can combine depth and velocity imaging using either two time-of-flight cameras or using the same device by alternating the modulation frequencies between successive video frames; color images can be obtained with a conventional camera.

Our technique offers a fundamentally new imaging modality that is ideally suited for fast motion. Optical flow applied to conventional RGB video is a complimentary technique: together, optical flow and D-ToF allow for the metric 3D velocity field to be estimated, which is otherwise not easily possible. In general, however, D-ToF is independent of the RGB flow and works robustly for cases where optical flow often fails, including untextured scenes and extremely high object velocities. We also discuss a mode for simultaneous range and velocity imaging. As with standard ToF imaging, our method requires a few subframes to be captured with different modulation signals. Using appropriate hardware (multi-sensor cameras or custom sensors with different patterns multiplexed into pixels of a single sensor), the method could be implemented as a true snapshot imaging approach. In our prototype system, we instead use rapid time-sequential acquisition of the required subframes, which is a common strategy for regular ToF imaging. In summary, this paper makes the following contributions:

- We introduce D-ToF as new modality of computational photography that enables instantaneous radial velocity estimation. Using multiple captures or implemented with multi-sensor setups, D-ToF records velocity, range, and color information.
- We derive a mathematical framework for velocity estimation with time-of-flight cameras, implement a prototype time-of-flight imaging system, and validate the proposed model extensively in simulation and with the prototype.
- We evaluate the imaging system using a range of different types of motion, for textured and untextured surfaces as well as indoors and under strong outdoor ambient illumination.
- We demonstrate that the velocities measured with our system can be combined with RGB flow, allowing for the metric 3D velocity field to be estimated on a per-pixel basis.

1.1 Limitations

As a fundamentally new imaging modality, D-ToF implemented with our experimental hardware has several limitations. Foremost, the resolution of the PMD sensor in our prototype is limited to 160×120 pixels and the signal-to-noise ratio (SNR) of the measured, Doppler-shifted signal is low. Together, these limitations result in low-resolution and noisy footage. We apply state-of-the-art denoising strategies which filter out most of the noise but sometimes result in "blobby" images. Further, D-ToF requires two frames to be acquired with different modulation frequencies. Currently, we capture these frames in sequence, which results in slight misalignment between the frames observed as velocity artifacts around depth discontinuities. However, there is a clear path to addressing all of these challenges: low-cost time-of-flight sensors providing QVGA or higher resolutions and significantly improved noise characteristics are already on the market. With access to signal control of illumination and on-sensor modulation, D-ToF could be readily implemented on high-quality consumer ToF cameras. Combining two synchronized ToF cameras with different frequencies would allow for misalignment artifacts to be mitigated.

Nevertheless, D-ToF shares other limitations with ToF, including the need for active illumination, limited range, and problematic processing in the presence of strong ambient illumination or objects with dark albedos, shadowed regions, and global illumination effects.

2 Related Work

Doppler-effect Measurements Since Christian Doppler discovered that the spectrum of astronomical objects shifts depending on their velocity [Doppler 1842], the Doppler effect has found widespread use in astronomical imaging, meteorology, traffic law enforcement, radiology, healthcare, and aviation. Doppler spectroscopy, for example, measures radial velocity of otherwise undetectable planets by observing wavelength shifts of their respective stars. The rate of expansion of the universe can be estimated by Doppler spectroscopy as well. Laser Doppler velocimetry is a common technique in healthcare, for example to measure blood flow. Usually, this technique uses two crossed, coherent laser beams to create a small volume of bright and dark fringe patterns; the rate of intensity fluctuation of particles moving through this volume gives rise to their velocity. Doppler radar is widely used in police speed guns, although gradually being replaced by lidar-based systems. Doppler lidar is also commonly used in many meteorological applications, such as wind velocity estimation. One common limitation of all Doppler measurements is that only movement along one particular direction, usually the line-of-sight, can be detected. All of these applications rely on the wave nature of light or sound, and therefore require coherent illumination or precise spectroscopic measurement apparatuses. We are the first to exploit incoherent, amplitude-modulated illumination and inexpensive time-of-flight (ToF) cameras for instantaneous imaging of both velocity and range. Our approach is a full-field imaging method, meaning that it *does not require the scene to be sequentially scanned* unlike most existing Doppler radar or lidar systems that only capture a single scene point at a time.

Computational Time-of-Flight Photography With consumer time-of-flight cameras such as Microsoft's Kinect One becoming widely available, research on computational time-of-flight imaging has become an emerging area throughout the last few years. New approaches to capture and visualize light transport have allowed physical lighting effects to be recorded and replayed [Velten et al. 2013; Heide et al. 2013] that the field of computer graphics has thus far only been able to simulate. Detailed analyses of temporal light transport in the frequency domain [Wu et al. 2012] or in the presence of global illumination [O'Toole et al. 2014; Gupta et al. 2014] have facilitated entirely new imaging modalities. For example, difficult inverse problems, such as non-line-of-sight imaging [Kirmani et al. 2009; Velten et al. 2012; Heide et al. 2014a], BRDF estimation [Naik et al. 2011], descattering [Heide et al. 2014b], and multipath separation [Kadambi et al. 2013], have become tractable.

In this manuscript, we analyze an effect not studied in prior work on computational time-of-flight imaging: the Doppler shift of objects in motion. We derive a mathematical framework and build a camera prototype implementing the described techniques; together, they allow us to optically encode object velocity into per-pixel measurements of modified time-of-flight cameras. By combining multiple cameras, we also demonstrate how to capture color, range, and velocity images simultaneously.

A technique loosely related to ours was recently proposed by Pandharkar et al. [2011]. Whereas they use a pulsed femtosecond illumination source to estimate motion of non-line-of-sight objects from differences in multiple captured images, we use the Doppler effect observed with conventional time-of-flight cameras within a single captured frame. In effect, their technique is related to optical flow methods that track features between successive video frames.

Notation	Description
$g(t)$	illumination signal at the light source
$s(t)$	illumination signal incident at the ToF sensor
$f_\psi(t)$	sensor reference signal
ω_g	illumination frequency
ω_f	on-sensor modulation frequency
ψ	programmable phase offset for sensor signal
ϕ	depth-dependent phase shift in illumination
$\Delta\omega$	Doppler frequency shift
$i_\psi(t')$	continuous, low-pass-filtered sensor image
$\boldsymbol{i}_\psi[t']$	discretely-sampled, low-pass-filtered sensor image

Table 1: *Notation table.*

Optical Flow in Computer Vision Optical flow [Horn and Schunck 1981; Barron et al. 1994] is a fundamental technique in computer vision that is vital for a wide range of applications, including tracking, segmentation, recognition, localization and mapping, video interpolation and manipulation, as well as defense. Optical flow from a single camera is restricted to estimating *lateral motion* whereas the Doppler is observed only for *radial motion* towards or away from the camera.

Wei et al. [2006] and Hontani et al. [2014] have demonstrated how to use correlation image sensors to estimate optical flow of fast motion. Although correlation image sensors are conceptually similar to ToF cameras, their methods are more similar in spirit to conventional optical flow by targeting lateral, rather than radial motion. In contrast to these methods, we analyze the Doppler effect of object motion to estimate per-pixel radial velocity without the need for optical flow. Lindner and Kolb [2009] as well as Hoegg et al. [2013] estimate lateral optical flow to compensate for object motion between the sequentially-captured ToF phase images from which depth is usually estimated. We can apply a similar strategy to mitigate alignment artifacts when subframes are captured sequentially, but the flow is not a core part of D-ToF.

3 Review of Time-of-Flight Imaging

In this section, we first review the conventional ToF image formation model for static scenes and then analyze how it behaves for objects in motion.

Time-of-flight cameras operate in continuous wave mode. That is, a light source illuminates the scene with an amplitude-modulated signal that changes periodically over time. Sinusoidal waves are often used in the ToF literature to approximate the true shape of the signals. Although we derive a full model for arbitrary periodic signals in the supplemental material, we restrict the derivation in this article to the sine wave model for simplicity of notation. Hence, the light source emits a temporal signal of the form

$$g(t) = g_1 \cos(\omega_g t) + g_0, \tag{1}$$

where ω_g is the illumination frequency. Assuming that the emitted light is reflected along a single, direct path by a stationary diffuse object at distance d, and that it is observed by a camera co-located with the light source, the signal reaching the camera is

$$\begin{aligned} s(t) &= s_1 \cos\left(\omega_s \left(t - \frac{2d}{c}\right)\right) + s_0 \\ &= s_1 \cos(\omega_s t + \phi) + s_0, \end{aligned} \tag{2}$$

with $s_0 = g_0 + b$, where b is the ambient illumination. In the case of a stationary scene, the frequency at the camera is the same as the illumination frequency: $\omega_s = \omega_g$. In Equation 2, the amplitude s_1 combines the illumination amplitude g_1, geometric factors such as 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 $\phi = -2d/c \cdot \omega_g$.

Theoretically, $s(t)$ could be directly sampled to estimate ϕ. However, illumination frequencies are usually in the order of tens to hundreds of MHz. Conventional solid state image sensors only provide sampling rates that are orders of magnitudes lower, and are hence inadequate for direct sampling of the phase. To overcome this limitation, time-of-flight camera pixels provide a crucial feature that makes them distinct from conventional camera pixels: before being digitally sampled, the incident signal is modulated by a high-frequency, periodic function $f_\psi(t)$ within each pixel. This on-sensor modulation is physically performed by an electric field that rapidly redirects incident photons-converted-to-electrons into one of two buckets within each pixel. The phase ψ and frequency ω_f of the modulation function are programmable. The general equation for the modulated signal is thus

$$\begin{aligned} \widetilde{i}_\psi(t) &= f_\psi(t) \cdot s(t) = \cos(\omega_f t + \psi) \cdot (s_1 \cos(\omega_s t + \phi) + s_0) \\ &= \frac{s_1}{2} \cos((\omega_f - \omega_s)t + \psi - \phi) + \\ &\quad \frac{s_1}{2} \cos((\omega_f + \omega_s)t + \psi + \phi) + s_0 \cos(\omega_f t + \psi). \end{aligned} \tag{3}$$

Usually, ToF cameras are operated in a *homodyne* mode where the illumination frequency and the reference frequency are identical: $\omega_f = \omega_g = \omega$. Under the common assumption of a stationary scene, we moreover get $\omega_s = \omega_g = \omega$, and Equation 3 simplifies to

$$\widetilde{i}_\psi(t) = \frac{s_1}{2} \cos(\psi - \phi) + \frac{s_1}{2} \cos(2\omega t + \phi + \psi) + s_0 \cos(\omega t + \psi). \tag{4}$$

To model the discretely sampled quantities measured by the sensor, we must account for a finite integration (exposure) time. The exposure time T of all cameras acts as a low-pass filter on the modulated signal before it is discretized by the sampling process of the sensor. Since the exposure time is usually significantly longer than the wavelength of the modulated signal $T \gg 1/\omega$, all frequency-dependent terms in Equation 4 vanish:

$$i_\psi(t') = \left(\widetilde{i}_\psi * \mathrm{rect}_T\right)(t') \approx \frac{s_1}{2} \cos(\psi - \phi). \tag{5}$$

The temporal low-pass filter $\mathrm{rect}_T(\cdot)$ is convolved with the incident signal — an operation that is analogous to the finite integration area of each sensor pixel in the spatial domain[2]. Finally, the modulated and low-pass-filtered signal is discretely sampled. Since Equation 5 is independent of the time of measurement t', depth and albedo can be robustly estimated.

To distinguish the continuous function $i_\psi(t')$ from its discretization, we denote the latter as $\boldsymbol{i}_\psi[t']$. For depth estimation, at least two measurements $\boldsymbol{i}_0[t']$ and $\boldsymbol{i}_{\pi/2}[t']$ are necessary that are usually recorded in quick succession, such that phase and depth can be estimated as

$$\boldsymbol{\phi}_{est}[t'] = \tan^{-1}\left(\frac{\boldsymbol{i}_{\pi/2}[t']}{\boldsymbol{i}_0[t']}\right), \text{ and } \boldsymbol{d}_{est}[t'] = \frac{c\,\boldsymbol{\phi}_{est}[t']}{2\omega}. \tag{6}$$

The same measurements are also used to estimate the albedo:

$$\boldsymbol{s_1}_{est}[t'] = \sqrt{(\boldsymbol{i}_0[t'])^2 + \left(\boldsymbol{i}_{\pi/2}[t']\right)^2}. \tag{7}$$

[2]In the optics community, the low-pass filter resulting from spatial sensor integration is known as the detector footprint modulation transfer function [Boreman 2001].

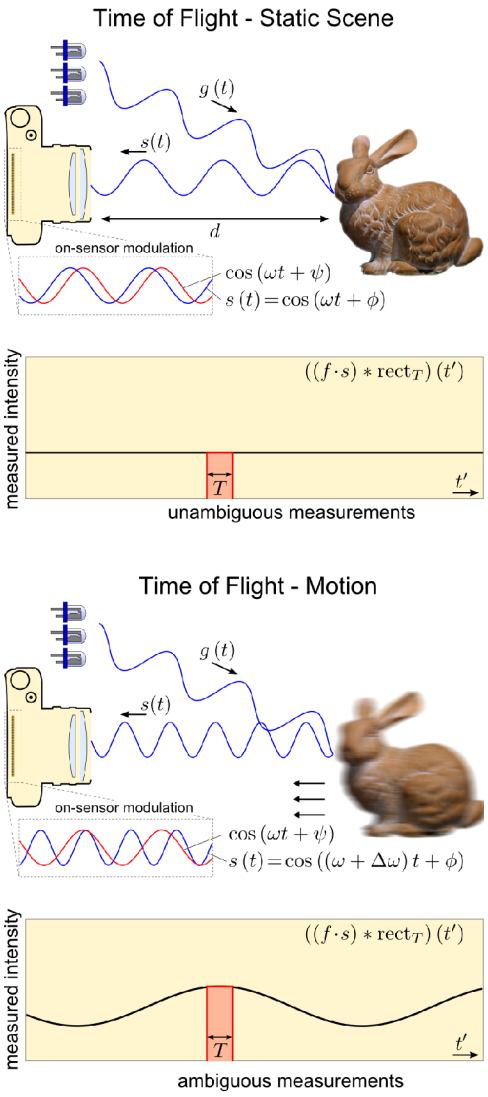

Figure 2: *Depth imaging. For static scenes, measurements are unambiguous: different phase shifts result in unique intensity measurements (top). For dynamic scenes, the Doppler shift results in a low-frequency beating pattern that makes measured intensities ambiguous, and hence prevents reliable depth estimation (bottom).*

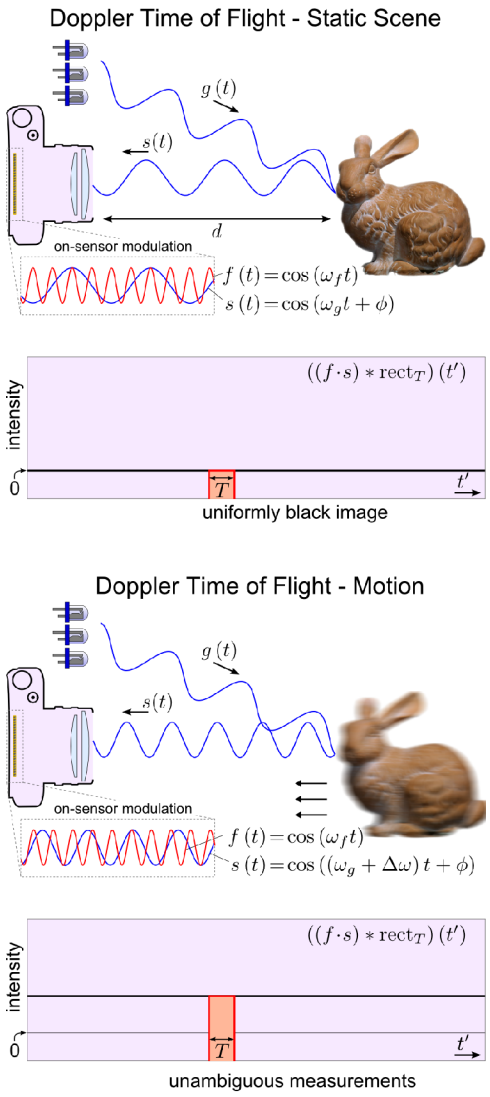

Figure 3: *Velocity imaging. Illumination ω_g and modulation ω_f frequencies are designed to be orthogonal within the exposure time T. For static scenes (top), this particular choice of frequencies will integrate to zero. The Doppler shift of moving scenes destroys the orthogonality and results in an approximately linear relationship between radial velocity and recorded intensity (bottom).*

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; Büttgen and Seitz 2008].

Time-of-Flight for Objects in Motion

The conventional time-of-flight image formation model breaks down when objects of interest move with a non-negligible radial velocity. In this case, the illumination frequency undergoes a Doppler shift [Doppler 1842] when reflected from an object in motion. The illumination arriving at the sensor is now frequency-shifted to $\omega_s = \omega_g + \Delta\omega$, where the change in temporal frequency $\Delta\omega$ depends on the radial object velocity as well as the illumination frequency:

$$\Delta\omega = \frac{v}{c}\,\omega_g. \tag{8}$$

For intuition, we consider the case of an approximately constant velocity v throughout the exposure time. If we continue to assume a homodyne setting with $\omega_f = \omega_g = \omega$, Equation 3 can be used to derive a new version of the low-pass-filtered sensor image (Eq. 5) for moving scenes:

$$i_\psi(t') \approx \frac{s_1}{2}\,\cos(-\Delta\omega t' + \psi - \phi). \tag{9}$$

Note that this equation is now dependent on the time of measurement. Unfortunately, the introduced temporal intensity variation makes it more difficult to estimate phase and therefore also depth. In audio signal processing, this time-dependent low-frequency artifact is known as a beating pattern. We illustrate it in Figure 2.

The phase estimate from Equation 6 is then distorted as

$$\phi_{est}[t'] = \tan^{-1}\left(\frac{i_{\pi/2}[t']}{i_0[t']}\right) + \Delta\omega t', \tag{10}$$

where the distortion $\Delta\omega t'$ linearly depends on the (unknown) object velocity. Note that, in practice, the estimated phase for moving objects corresponds to its average throughout the exposure.

To summarize, in the homodyne setup, where the frequency of the light source and the frequency of the camera reference signal are identical, the Doppler shift introduced by moving objects results in mismatched frequencies on the image sensor. This situation is closely related to *hetereodyne* time-of-flight imaging (e.g., [Dorrington et al. 2007]), which generalizes the conventional homodyne capture mode to arbitrary combinations of illumination and sensor modulation frequencies. For static scenes, the heterodyne imaging mode (e.g., [Conroy et al. 2009]) may be beneficial in certain situations, but a major limitation of heterodyne ToF is that multiple (>2) measurements have to be captured to reliably estimate phase and depth. Since the beating pattern is usually of very low frequency, a significant amount of time needs to pass between the two measurements for reliable phase estimation. For moving objects, the necessity to capture multiple images would place severe constraints on the velocity. To facilitate reliable velocity estimation, we derive a new computational time-of-flight imaging methodology in the following section. Inspired by the general concept of orthogonal frequency-division multiplexing (OFDM, e.g. [Li and Stuber 2006]), D-ToF uses illumination and on-sensor modulation frequencies that are orthogonal within the exposure time of the camera. Using this choice of frequencies along with a newly-devised reconstruction method, we demonstrate the first approach to per-pixel radial velocity estimation.

4 Doppler-based Velocity Imaging

As illustrated in Figure 2 (bottom), the low-frequency beating pattern created by the Doppler effect makes it difficult or impossible to capture reliable Doppler frequency and phase information. Consider the following example: a road cyclist travels at a speed of $v = 10\frac{m}{s}$ towards the camera. For an illumination frequency of 50 MHz (i.e. $\omega_g = 50 \cdot 10^6 \cdot 2\pi/s$), the observed Doppler shift is only

$$\Delta\omega = \frac{v}{c}\,\omega_g = \frac{10\frac{m}{s}}{300 \cdot 10^6 \frac{m}{s}} \cdot 50 \cdot 10^6 \frac{2\pi}{s} \approx 1.67\frac{2\pi}{s} \qquad (11)$$

A frequency shift of only 1.67 Hz may seem small enough to be safely ignored. However, we show in the following that even such a minute change contains valuable information that can be used for velocity estimation.

4.1 Velocity Imaging via Orthogonal Frequencies

Inspired by multiplexing techniques in digital communication, we devise an unconventional way to extract velocity information from the small Doppler shift observed by a ToF camera. We can interpret the camera system as a communication channel and consider the illumination a carrier signal. The carrier is optically modified by moving objects — we observe a change in carrier amplitude, phase, and frequency. The secondary modulation in the sensor followed by a low-pass filter of the exposure time corresponds to the demodulation process in communication. Conventional communication channels use orthogonal frequencies; any inter-carrier interference (which could be caused by a frequency drift) is a polluting signal (see e.g. [Li and Stuber 2006]). For Doppler ToF, we deliberately design the frequencies in the receiver and transmitter to be orthogonal, such that the (usually polluting) inter-carrier interference carries the desired velocity information.

For the application of direct velocity imaging, we would like to ensure that the measured signal for a stationary object is zero (or a constant intensity offset). We can achieve this by operating the ToF camera in heterodyne mode with two orthogonal frequencies ω_g and ω_f. While any two sine waves with frequencies $\omega_g \neq \omega_f$ will be orthogonal for sufficiently long integration times, this is not the case for finite integrals (exposures) in the presence of low frequency beating patterns. Designing both frequencies to be orthogonal is done by setting

$$\omega_g = k\,\frac{2\pi}{T} \quad \text{and} \quad \omega_f = l\,\frac{2\pi}{T} \quad \text{with}\, k, l \in \mathbb{N}, k \neq l, \qquad (12)$$

i.e. having the exposure time T be an integer multiple of the period of both signals. It is then easy to show from Equation 3 that

$$i_\psi = \int_0^T \widetilde{i}_\psi(t)\,dt = 0 \qquad (13)$$

for stationary objects ($\omega_s = \omega_g$). In practice, we set $l = k + 1$ and we set $k = \omega_g T/2\pi$, which depends on T and the desired frequency ω_g.

Given these two orthogonal frequencies we now use the inter-carrier interference to extract valuable information about the Doppler shift. We achieve this by computing the ratio of a heterodyne measurement and a homodyne measurement. Using only the low-frequency terms from Equation 3, this ratio can be expressed as[3]:

$$\begin{aligned}
r =& \frac{\int_0^T \cos(\omega_f t + \psi) \cdot (s_1\,\cos((\omega_g + \Delta\omega)t + \phi) + s_0)\;dt}{\int_0^T \cos(\omega_g t + \psi) \cdot (s_1\,\cos((\omega_g + \Delta\omega)t + \phi) + s_0)\;dt} \\
\approx& \frac{\int_0^T \frac{s_1}{2}\,\cos((\omega_f - \omega_g - \Delta\omega)t + \psi - \phi)\,dt}{\int_0^T \frac{s_1}{2}\,\cos(-\Delta\omega t + \psi - \phi)\,dt} \\
=& \frac{\frac{s_1}{2(\omega_f - \omega_g - \Delta\omega)}\,[\sin((\omega_f - \omega_g)t - \Delta\omega t + \psi - \phi)]_0^T}{\frac{s_1}{-2\Delta\omega}\,[-\Delta\omega t + \psi - \phi)]_0^T} \\
=& \frac{-\Delta\omega}{\omega_f - \omega_g - \Delta\omega} \\
& \cdot \underbrace{\frac{\sin((\omega_f - \omega_g)T - \Delta\omega T + \psi - \phi) - \sin(\psi - \phi)}{\sin(-\Delta\omega T + \psi - \phi) - \sin(\psi - \phi)}}_{=1} \\
\approx& \frac{-\Delta\omega}{\omega_f - \omega_g}
\end{aligned} \qquad (14)$$

since $(\omega_f - \omega_g)T = (k - l)\,2\pi$, and $\Delta\omega \ll \omega_f - \omega_g$.

Figure 4 shows the model derived here. On the left side, we see the full model without any approximations (i.e. without neglecting high frequency components in Eq. 14). Although the image formation is nonlinear, for a relative large range of metric velocities it is very well approximated (Fig. 4, center left) by our linear model (Eq. 14). We experimentally verify the model using our camera prototype (Fig. 4, right). With known, orthogonal illumination and modulation frequencies ω_g, ω_f, it is therefore straightforward to compute the Doppler shift $\Delta\omega$ from Equation 14. The ratio image r can be interpreted as a direct measurement of the instantaneous per-pixel radial velocity.

We note that this approach still requires two measurements: one heterodyne image and one homodyne image. There are several possible solutions for either acquiring these truly simultaneously, or they can be acquired in quick succession. For instantaneous measurements, two synchronized ToF sensors can be mounted in a co-axial setup; one of the sensors is modulated with the same

[3] Without loss of generality, we assume an exposure interval of $[0 \ldots T]$

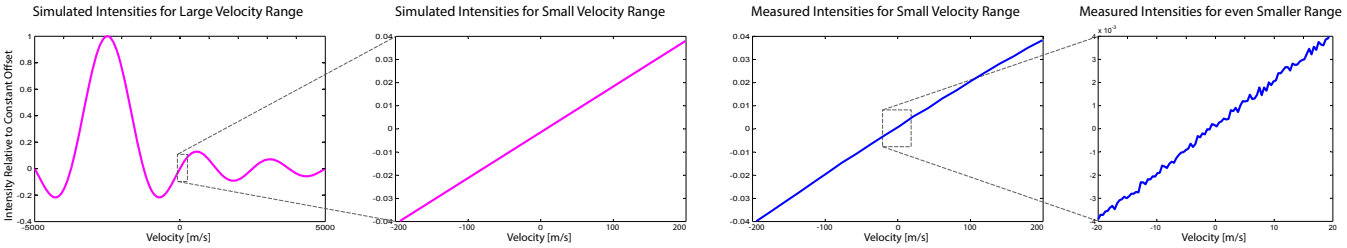

Figure 4: *Simulated and measured intensities for a range of different velocities. Although the mapping from radial velocity to measured intensity is generally nonlinear (left), throughout a large range of velocities the conversion is approximately linear (center left). We verify the predicted mapping using our prototype camera (right). These particular measurements were captured with a static scene, and acquired with a modulation frequency of $\omega_f = 60\,MHz$ and an illumination frequency of $\omega_g = 60\,MHz + 1\,KHz + \Delta\omega$. Thus, the Doppler shift for an object moving at a specific velocity was programmed into the illumination frequency for this particular experiment.*

frequency as the light source (ω_g), while the other uses a slightly different frequency $\omega_f \neq \omega_g$. This approach is similar in spirit to multi-sensor HDR imaging [Tocci et al. 2011].

Instead of using two distinct sensors, it would also be possible to multiplex pixels with two different modulation frequencies onto the same image sensor, either in alternating scanlines or in a checkerboard pattern. Again, this concept is similar in spirit to techniques that have been proposed for HDR cameras [Yasuma et al. 2010; Gu et al. 2010].

A third possibility is to rapidly alternate between two modulation frequencies using a single ToF camera. In this case, the measurements are not truly instantaneous, and alignment problems may occur for very fast motions. However, the two measurements can be taken immediately after each other, as fast as the camera hardware allows, e.g. at 30 or 60 Hz. We follow this approach as it only requires a single ToF camera. Note that, similar to heterodyne depth estimation [Dorrington et al. 2007], the Doppler shift can also be estimated directly from the low-frequency beating pattern, but at the cost of requiring multiple measurements that are much more widely spaced in time (hence not suitable for velocity estimation).

Finally, we note that the model from Equation 14 only holds for sinusoidal modulation functions. If other periodic signals are being used, additional harmonic frequency components are introduced, which distort the measurements for both stationary and moving targets. However, these offsets are systematic and can be calibrated for a specific ToF camera/lights source combination (see Section 5, and the supplemental material).

4.2 Simultaneous Range and Velocity

In many applications it may be useful to obtain both velocity and range measurements at the same time. As in standard ToF imaging, this can be achieved by capturing a second homodyne measurement with the phase offset by $\pi/2$. Simultaneous range and velocity imaging therefore requires a total of three measurements: a heterodyne image with $\psi = 0$, a homodyne image with $\psi = 0$, and a homodyne image with $\psi = \pi/2$.

As discussed in Section 3, motion introduces a velocity-dependent distortion $\Delta\omega t'$ of the depth measurement (Eq. 10). However, since the distortion linearly depends on the Doppler shift $\Delta\omega$, which is known from the velocity estimation step (Eq. 14), we can now correctly estimate the phase delay (and hence the depth) from Equation 10. This only requires an additional calibration step to obtain $\Delta\omega t'$ for a specific velocity, which corresponds to estimating the time offset t' between the start of the exposure time and the reference time for signal generation in the camera and light source.

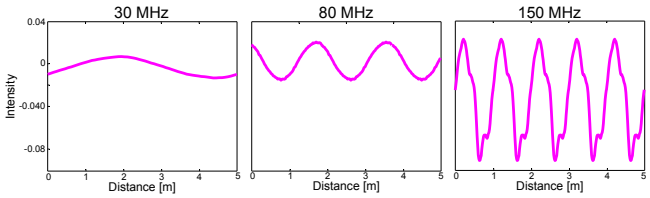

Figure 5: *Depth-dependent offset introduced by higher-order frequency components for a range of modulation frequencies. These offsets are calibrated in a one-time offline process and then used to correct the raw phase measurements on a per-pixel basis.*

As mentioned, simultaneous velocity and range imaging requires three distinct measurements. We note that the illumination signal is the same for all three measurements, only the reference signal for the camera changes. As in the case of velocity-only imaging, this means that all three measurements can potentially be acquired at the same time using either multiple sensors with a shared optical axis, or a special sensor design with interleaved pixels. If neither option is available, rapid frame-sequential imaging is also possible.

5 Implementation

Hardware For all physical experiments, we use an experimental time-of-flight camera system that comprises a custom RF modulated light source and a demodulation camera based on the PMD Technologies PhotonICs 19k-S3 sensor (see Fig. 1). The light source is an array of 650 nm laser diodes driven by iC-Haus constant current driver chips, type ic-HG. We use a PMD CamBoard nano development kit with a clear glass sensor that has the near IR bandpass filter removed, in combination with an external 2-channel signal generator to modulate the sensor and synchronize the light source. Our setup is similar to commercially-available time-of-flight cameras and the proposed algorithms could be easily implemented on those. Unfortunately, developers usually do not have access to illumination and modulation frequencies of these devices, requiring the construction of custom research prototype cameras. The maximum illumination and demodulation frequency of our prototype is 150 MHz, but we run all of the presented results with 30 MHz. The modulation signals are nearly sinusoidal, but contain multiple low-amplitude harmonic components. To avoid systematic errors in depth and velocity estimation, these components must be calibrated as described in the following.

Correcting for Higher-order Harmonics Our camera prototype has the drawback that the periodic modulation functions are not

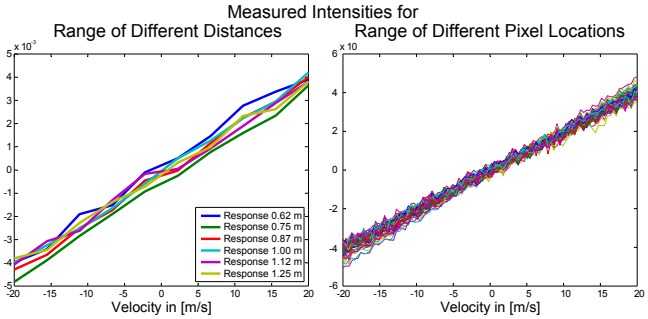

Figure 6: *Experimental verification of the imaging system for varying object velocities and depths (left) as well as velocity-dependent behavior for a range of different pixel locations on the sensor (right). All of this data is captured using a large planar target perpendicular to the camera and sweeping the illumination frequency (to simulate different Doppler shifts) and phase (to simulate different object distances).*

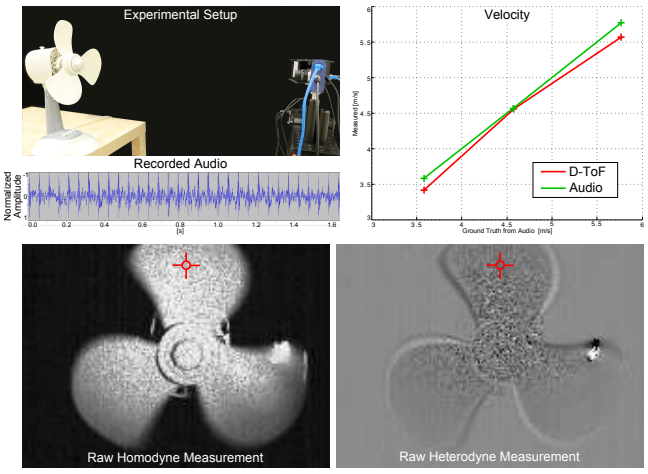

Figure 7: *Experimental validation of velocity estimation using a fan with adjustable rotation speed (three settings). We measure the ground truth velocity of the rotating blades (top left) by analyzing audio recordings (top, lower left). The top right plot shows the velocity measured by D-ToF compared to the ground truth for a varying rotation speed. As the speed becomes larger, estimation errors increase to a maximum of about 0.2 m/s. The bottom row shows the unprocessed full-field measurements of the homodyne (left) and the heterodyne (right) frequency setting with the pixel indicated for which we plotted the velocities on the top right.*

perfectly sinusoidal, although they are very close. In addition to the fundamental frequency, this introduces higher-order harmonic components to the modulation signal. Please refer to the supplemental material for a detailed derivation of the image formation in these conditions. Unfortunately, the higher-order components are generally not orthogonal, thus they can cause a phase-dependent offset. We calibrate this offset for different modulation frequencies and phase shifts ψ using a static target. The depth-dependent offsets are plotted for different modulation frequencies in Figure 5.

This offset is calibrated in an offline process and raw phase measurements can be corrected digitally using a lookup table. Note that for relatively low modulation frequencies, such as 30 MHz, we find a fairly large depth range (around 1 m) to be almost independent of this offset. In practice, it is therefore relatively easy to remove the higher-order frequency components.

Calibrating Phase Response As is standard practice in time-of-flight cameras, we calibrate the physical intensity response for different phase shifts ϕ in an offline calibration. Following [Lindner and Kolb 2006], we measure the physical intensity response for a phase sweep of the illumination frequency and fit a fifth-order polynomial to the measurements. This is used as a lookup table for converting phase to depth rather than solving Equation 6 directly. With our prototype, we measure a notable zeroth-order component of the fitted polynomial, corresponding to fixed pattern phase noise. This is easily corrected with the lookup table. Any other illumination-specific terms, for example introduced by the baseline between camera and light source, are automatically calibrated with the described procedure and do not require additional processing.

Verification of Calibration Procedure The two calibration procedures described above are performed for all spatial locations on the sensor independently. To verify our calibration routines, we image a static target and apply a frequency and phase sweep to the modulation function, simulating objects at different velocities and depths. The results shown in Figure 4 (left) demonstrate that the measured intensities for a constant phase but varying Doppler shift follow the model derived in the Section 4. Other than a small amount of noise, which is mostly due to a relatively low signal-to-noise ratio, the curve is linear and behaves as predicted. In Figure 6 (left), we verify experimental measurements for a range of different phase offsets in the modulation frequency. This simulates objects at various depths, as indicated in the legend. Finally, we also test the velocity-dependent behavior for a range of different pixels over the sensor location and show results in Figure 6 (right). The remaining variance over pixel locations and phases is minimal.

Figure 7 shows another experiment that we used to verify the accuracy of our prototype D-ToF camera. In this example, we adjusted the speed of a rotating fan and imaged its blades such that, throughout the time it takes for a single blade to move across a pixel, forward motion is observed by that pixel. The exposure time of the ToF camera was set to 1.5 ms and the fan was captured from a frontal perspective (raw homodyne and heterodyne measurements shown in Fig. 7 bottom). We manually measured the slope of the fan blades, which is constant over the entire blades. The radius of the plotted position was measured, allowing us to calculate the "ground truth" velocity when the rotation speed of the fan is known. Since the exact rotation speed is not actually known, we measure it by mounting a small pin on one of the blades and mounting a piece of flexible plastic in front of the fan, such that the rotating pin strikes the plastic exactly once per revolution, creating a distinct sound. We record the sound (sampled at 44 KHz) of this setup to estimate the ground truth velocity of the fan blades, observed by one pixel, which is compared with the corresponding D-ToF estimate (Fig. 7, top right). For this experiment, the estimation error is always below 0.2 m/s. Errors are mainly due to the low SNR of the measured Doppler-shifted signal.

Subframe Alignment Although the required heterodyne and homodyne shots could be captured simultaneously using multi-sensor configurations, they have to be captured in an alternating fashion using the single-sensor solution used in this paper. Since we are dealing with moving objects, the individual shots cannot be assumed to be perfectly aligned, which results in velocity artifacts around edges in the scene. We can mitigate, although not completely remove, these artifacts by computing a SIFT flow on the raw data and warping them to a reference frame. While not perfect, the SIFT flow delivered sufficiently good warps for most captures.

Figure 8: *Complex scene with ambient illumination and a large depth range. The velocity is robustly estimated within the range of the illumination (approx. 5m inside), even in outdoor settings.*

Denoising With our system, we capture an extremely small frequency shift (in the Hz range) relative to the modulation frequency (the MHz range). Additionally, the quantum efficiency of emerging time-of-flight sensors is still far from that of modern solid state sensors [Erz and Jähne 2009]. Therefore, the slight Doppler shift in our prototype is strongly affected by Poisson noise. Standard denoising methods fail in strong Poisson noise scenarios. We apply a binning-based non-local means denoising strategy to all captured velocity maps (see Fig. 9). Please see the supplement for more details and denoising comparisons.

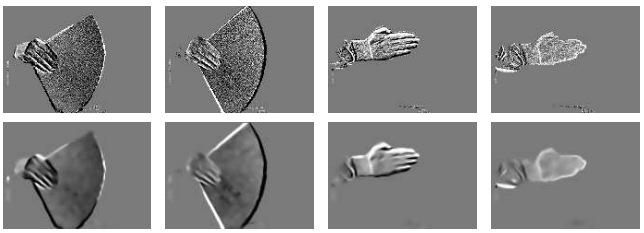

Figure 9: *Velocity maps color-coded in grayscale. The maps computed from raw measurements (top) are corrupted by Poisson noise. To account for this, we apply a binning-based non-local means-type denoiser to the reconstructed velocity images (bottom).*

6 Experimental Results

We show results captured with our prototype imaging system in Figures 1, 8, 10, 11, 12, 13, 14, and in the supplement. The results validate the proposed imaging system for a variety of challenging indoor and outdoor scenes. Color images are recorded with the same exposure time as the time-of-flight camera. Most of the scenes have a slight red tint, because we work with eye-safe red illumination in the visible spectrum. Like current commercial ToF cameras, future implementations of this system would most likely use invisible, near infrared wavelengths to encode velocity and depth information. The reconstructed velocity maps are color-coded; absolute units are indicated in the color bars. As expected, static scenes result in a constant velocity map whereas velocity is directly encoded in the measurements and subsequently reconstructed for each sensor pixel independently. In addition to the velocity maps, Figures 1, 11, 13, 14 also show the corresponding depth maps that can be estimated from an additional capture as well as the velocity maps (see Sec. 4.2).

The selection of scenes shows a wide range of motion types that

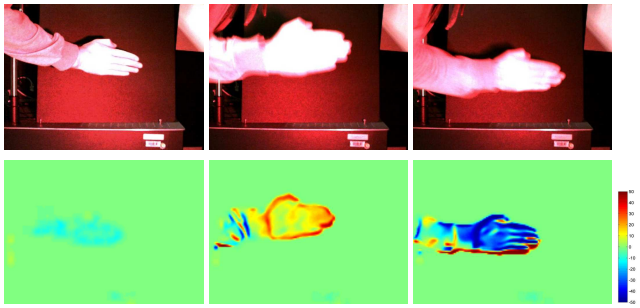

Figure 10: *This result shows a periodic motion of a hand along the optical axis. The static scene on the left results in no response of the sensor, whereas forward (center) and backward (right) movement result in positive and negative responses, respectively.*

can be reconstructed with the proposed method, but it also highlights several challenges of D-ToF and ToF in general. D-ToF requires two frames to be captured, and they must be aligned if recorded with a single camera. In some instances, such as Figures 10 and 12, the alignment is challenging and any errors will propagate into the velocity maps, especially around depth-discontinuities. These artifacts could be mitigated by optimizing the camera firmware to minimizing switching time between the subframes or by using two co-axial ToF cameras. Objects with dark albedos, as for example observed in Figure 11, are challenging for any ToF method because only a small amount of the coded illumination is reflected back to the camera. Similarly, shadows are very challenging and often result in either no depth/velocity estimation or errors (sweater in Fig. 8 and regions between fingers in Fig. 13). Whereas some of these limitations can be overcome with better hardware, others are inherent to the time-of-flight approach. Please see Section 8 for a more detailed discussion and the supplemental video for more results.

7 Towards the 3D Velocity Field

Optical flow computed from conventional video sequences estimates the 2D projection of the 3D flow field onto the image plane. The radial component is usually lost. Furthermore, optical flow is an ill-posed problem and may fail in many scenarios. Doppler ToF addresses two problems of optical flow: first, it can help in cases where optical flow fails either due to large displacements or missing scene structures. Second, our technique also helps in cases

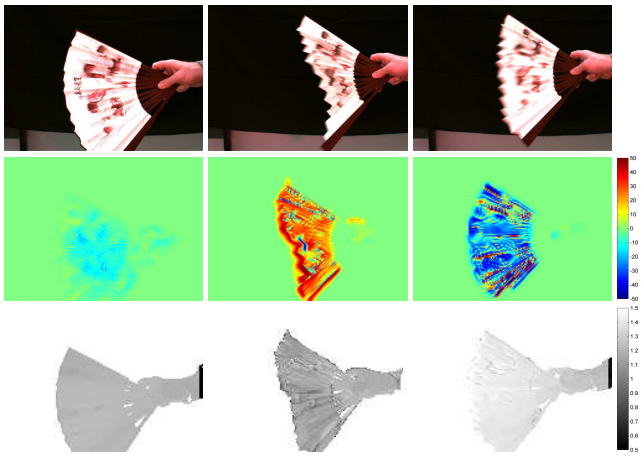

Figure 11: *This result shows periodic motions in z for a textured object. Although the estimated velocity is mostly correct, shadows and dark scene parts are challenging for robust velocity estimation.*

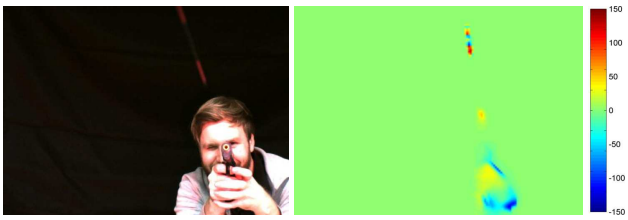

Figure 12: *Even extremely fast motion, such as a bullets shot with a spring airsoft gun, can be captured with our system. The airsoft gun is being advertised as shooting bullets with 99 m/s; we measure a radial velocity of 98.2 m/s (average of the peak pixels).*

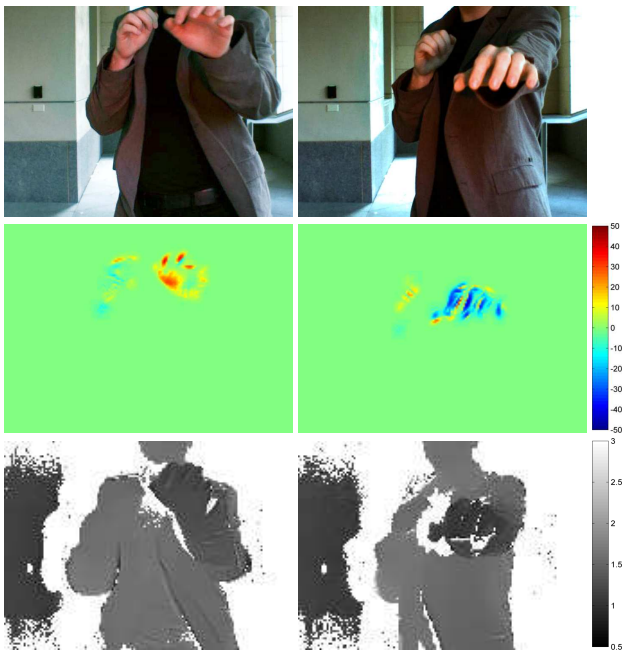

Figure 13: *We envision a wide range of applications for our technique, including gaming and human-computer interaction.*

where the optical flow estimation is successful; in this case, we can

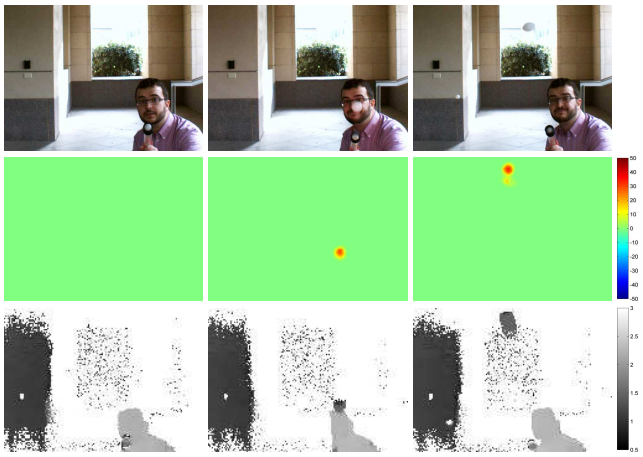

Figure 14: *Physical props for gaming, such as ping pong balls fired with this toy gun, could be tracked and enable new HCI techniques.*

recover the 3D metric flow by combining metric radial velocity and the 2D optical pixel flow.

Figure 15 shows an example scene where regular optical flow [Liu 2009] as well as SIFT-flow [Liu et al. 2008] fail due to limited structure in the scene. Our method can successfully capture the velocity of the objects and could also lead to a proper segmentation of the scene. Note that having additional depth estimates for conventional flow would also only be of limited help since flat surfaces also do not deliver enough features for correspondence matching.

Figure 16 shows a scene where the optical flow estimate is reasonable. In this case, the orthogonal component that our method captures completes the 2D spatial flow estimates and uniquely determines the full metric 3D flow. Given the optical flow estimates f_x, f_y for the horizontal and vertical image coordinates, one can compute the metric velocity vectors $v_x = \frac{f_x \cdot Z}{F}$, $f_y = \frac{f_x \cdot Z}{F}$, where F is the focal length of the lens and Z the corresponding depth estimate (see [Honegger et al. 2013]). In conjunction with the velocity estimate v_z in the orthogonal direction along the optical axis, the full 3D metric flow is $V = (v_x, v_y, v_z)$. An example is shown in Figure 16. Please note that the 3D flow field is only as reliable as the estimated radial velocity and the RGB 2D flow. If one of them fails, so will the 3D flow.

8 Discussion

In summary, we propose a new computational imaging modality that directly captures radial object velocity via Doppler Time-of-Flight Imaging. We demonstrate a variety of experimental results captured with a prototype camera system for different types of motions and outdoor settings. The mathematical models are extensively validated in simulation and experiment. We also show the optional combination of footage captured using an RGB camera with the depth and velocity output of our coded time-of-flight camera. Together, this data represents simultaneous per-pixel RGB, depth, and velocity estimates of a scene and allows for the 3D velocity field to be estimated. We envision applications in a wide range of computer vision problems, including segmentation, recognition, tracking, super-resolution, spatially-varying motion deblurring, and navigation of autonomous vehicles.

D-ToF is complimentary to optical flow. It allows for the depth bias of xz-flow to be removed and enables recording of the metric 3D velocity field of the scene. However, if only radial velocity is

Figure 15: *Failure case of optical flow for a moving, but untextured scene (left). Optical flow [Liu 2009] and SIFT flow [Liu et al. 2008] for two succeeding color frames are shown in the second and third column; the 2D flow vectors are color-coded with the shown color wheel (insets). Both methods cannot recover the true 2D motion of the fan and wrongly segment the scene. Our orthogonal velocity estimate can resolve this problem and properly segment the scene.*

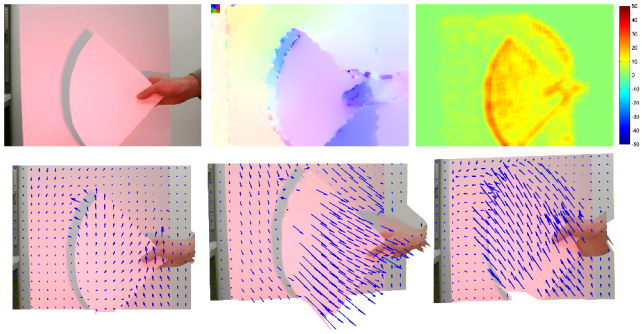

Figure 16: *Towards 3D flow: when optical flow succeeds, the full 3D metric flow is uniquely estimated from both 2D pixel flow and the radial velocity maps. The top images show a frame where optical flow computed reasonable estimates. The bottom shows full 3D velocity estimate for different views. Note that the optical flow helps us to determine that fan's velocity is slightly rotated to the upper right, where the center of rotation is located (bottom left).*

required, our technique can also be used stand-alone, independent of optical flow.

Limitations and Future Work Commercially available ToF sensors today are low-resolution and their quantum efficiency and noise characteristics are not comparable with modern CMOS sensors. We expect future generations of ToF sensors to deliver significantly higher image quality, which would directly benefit D-ToF as well. Higher modulation frequencies would directly improve the signal-to-noise ratio in our setup, because the Doppler effect is proportional to these frequencies. For eye-safe operation, we use diffused laser diodes that operate in the visible spectrum in combination with a ToF sensor that has its visible spectrum cutoff filter removed. The laser illumination is therefore visible in all of the RGB images as a red tint. Future implementations of our system would operate the time-of-flight camera in the near infrared spectrum, as is common practice in commercial ToF cameras. Finally, all presented techniques could easily be implemented on consumer time-of-flight cameras with the appropriate level of access to the system firmware or driver software. We hope that vendors will give researchers and developers the opportunity to modify modulation and illumination frequencies of their devices in the near future. With access to these frequencies, we could port the proposed techniques to available consumer devices.

Conclusion Time-of-flight cameras have entered the consumer market only a few years ago, but transformed the way machines perceive the world. Human-computer interaction, medical imaging, robotics and machine vision, navigation for self-driving cars and quadcopters, and many other fundamental computer vision tasks have seen dramatic improvements using these devices. With Doppler Time-of-Flight, we hope to contribute a fundamentally new imaging modality that will have an impact on all of these applications. The possibility of implementing the proposed techniques on existing consumer devices makes Doppler Time-of-Flight a particularly attractive computational photography technique.

9 Appendix

In this appendix, we derive expressions for the proposed imaging systems using phasor notation. These may be more intuitive for some readers, especially those familiar with the communication literature. In particular, we assume orthogonal illumination and modulation frequencies and derive the measured intensity of a dynamic scene (see Eqs. 13, 14) as

$$\begin{aligned} i_\psi &= \int_0^T \cos\left(2\pi\frac{k}{T}t+\theta\right) a\cos\left(2\pi\frac{l}{T}t+\phi\right)\,\mathrm{d}t \\ &= \frac{1}{2}\mathcal{R}\left\{\int_0^T \underbrace{e^{i\theta}}_{\tilde{A}} e^{i2\pi\frac{k}{T}t}\left(\underbrace{ae^{i\phi}}_{\tilde{B}} e^{i2\pi\frac{l}{T}t}\right)^*\,\mathrm{d}t\right\} \\ &= \frac{1}{2}\mathcal{R}\left\{\int_0^T \tilde{A}e^{i2\pi\frac{k}{T}t}\left(\tilde{B}e^{i2\pi\frac{l}{T}t}\right)^*\,\mathrm{d}t\right\} \\ &= \frac{1}{2}\mathcal{R}\left\{\tilde{A}\tilde{B}^*\int_0^T e^{i2\pi\frac{k-l}{T}t}\,\mathrm{d}t\right\} = \frac{1}{2}\mathcal{R}\{\tilde{A}\tilde{B}^*\}T\delta_{kl} \end{aligned} \tag{15}$$

$\tilde{A}, \tilde{B}$ are the complex phasor amplitudes containing the amplitude and phase dependency. δ_{kl} is the Kronecker delta; for perfectly static scenes, this expression is zero. Note that phasors can usually not be multiplied. By using the complex conjugate, this is possible while implicitly assuming that high frequency can be ignored [Ceperley 2015].

Assuming that objects are moving, a Doppler shift δ/T is introduced. Frequency shifts like this have been analyzed for OFDM as inter carrier interference. We get the following complex interference pattern I_δ:

$$\begin{aligned} I_\delta &= \int_0^T e^{i2\pi\frac{k}{T}t}\left(e^{i2\pi\frac{m-\delta}{T}t}\right)^*\,\mathrm{d}t = \int_0^T e^{i2\pi\frac{k-m+\delta}{T}t}\,\mathrm{d}t \\ &= \frac{T}{i2\pi(k-m+\delta)}\left(e^{i2\pi(k-m+\delta)}-1\right) \end{aligned} \tag{16}$$

Acknowledgements

The authors would like to thank Lei Xiao, Refael Whyte, Adrian Dorrington, Tom Malzbender, Charles Fraccia, Achuta Kadambi, and the anonymous reviewers for valuable feedback and help with the experiments. Felix Heide was supported by a Four-year Fellowship from the University of British Columbia. Matthias Hullin was supported by the X-Rite Chair for Digital Material Appearance. Gordon Wetzstein was supported by a Terman Faculty Fellowship. Wolfgang Heidrich was supported by Baseline Funding of the King Abdullah University of Science and Technology and a NSERC Discovery Grant.

References

BARRON, J., FLEET, D., AND BEAUCHEMIN, S. 1994. Performance of optical flow techniques. *IJCV 12*, 1, 43–77.

BOREMAN, G. D. 2001. *Modulation Transfer Function in Optical and ElectroOptical Systems*. SPIE Publications.

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.

CEPERLEY, P., 2015. Resonances, waves and fields. `http://resonanceswavesandfields.blogspot.com/2011/04/28-valid-method-of-multiplying-two.html`. [Online; accessed 20-January-2015].

CONROY, R., DORRINGTON, A., KUNNEMEYER, R., AND CREE, M. 2009. Range Imager Performance Comparison in Homodyne and Heterodyne Operating Modes. In *Proc. SPIE 7239*.

DOPPLER, C. J. 1842. Über das farbige Licht der Doppelsterne und einiger anderer Gestirne des Himmels. *Abhandlungen der Königl. Böhm. Gesellschaft der Wissenschaften 12*, 2, 465–482.

DORRINGTON, A. A., CREE, M. J., PAYNE, A. D., CONROY, R. M., AND CARNEGIE, D. A. 2007. Achieving sub-millimetre precision with a solid-state full-field heterodyning range imaging camera. In *Proc. Meas. Sci. Technol.*, vol. 18.

ERZ, M., AND JÄHNE, B. 2009. Radiometric and spectrometric calibrations, and distance noise measurement of ToF cameras. In *Dynamic 3D Imaging*. Springer, 28–41.

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

GU, J., HITOMI, Y., MITSUNAGA, T., AND NAYAR, S. 2010. Coded Rolling Shutter Photography: Flexible Space-Time Sampling. In *Proc. ICCP*.

GUPTA, M., NAYAR, S. K., HULLIN, M., AND MARTIN, J. 2014. Phasor Imaging: A Generalization Of Correlation-Based Time-of-Flight Imaging. Tech. rep., Jun.

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. *OSA Opt. Exp. 22*, 21, 26338–26350.

HOEGG, T., LEFLOCH, D., AND KOLB, A. 2013. Real-time Motion Compensation for PMD-ToF Images. In *Lecture Notes in Computer Science*, vol. 8200.

HONEGGER, D., MEIER, L., TANSKANEN, P., AND POLLEFEYS, M. 2013. An open source and open hardware embedded metric optical flow CMOS camera for indoor and outdoor applications. In *Proc. ICRA*, IEEE, 1736–1741.

HONTANI, H., OISHI, G., AND KITAGAWA, T. 2014. Local estimation of high velocity optical flow with correlation image sensor. In *Proc. ECCV*, 235–249.

HORN, B., AND SCHUNCK, B. 1981. Determining optical flow. *Artificial Intelligence 17*, 185–203.

KADAMBI, A., WHYTE, R., BHANDARI, A., STREETER, L., BARSI, 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.

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.

LI, Y., AND STUBER, G. 2006. *Orthogonal Frequency Division Multiplexing for Wireless Communications*. Springer.

LINDNER, M., AND KOLB, A. 2006. Lateral and depth calibration of PMD-distance sensors. In *Advances in Visual Computing*. Springer, 524–533.

LINDNER, M., AND KOLB, A. 2009. Compensation of Motion Artifacts for Time-of-Flight Cameras. In *Proc. Dynamic 3D Imaging*. 16–27.

LIU, C., YUEN, J., TORRALBA, A., SIVIC, J., AND FREEMAN, W. T. 2008. SIFT flow: Dense correspondence across different scenes. In *Computer Vision–ECCV 2008*. Springer, 28–42.

LIU, C. 2009. *Beyond pixels: exploring new representations and applications for motion analysis*. PhD thesis, MIT.

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.

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.

PANDHARKAR, R., VELTEN, A., BARDAGJY, A., LAWSON, E., BAWENDI, M., AND RASKAR, R. 2011. Estimating motion and size of moving non-line-of-sight objects in cluttered environments. In *Proc. CVPR*, 265–272.

TOCCI, M., KISER, C., TOCCI, N., AND SEN, P. 2011. A versatile HDR video production system. *ACM Trans. Graph. (SIGGRAPH) 30*, 4, 41.

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., BARSI, 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.

WEI, D., MASUREL, P., KURIHARA, T., AND ANDO, S. 2006. Optical flow determination with complex-sinusoidally modulated imaging. In *Proc. ICSP*, vol. 2.

WU, D., WETZSTEIN, G., BARSI, 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*, 542–555.

YASUMA, F., MITSUNAGA, T., ISO, D., AND NAYAR, S. K. 2010. Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum. *IEEE TIP 19*, 9, 2241–2253.