

Signal Processing for Time-of-Flight Imaging Sensors

Ayush Bhandari and Ramesh Raskar

CONTENTS

I	Introduction	2
I-A	A Brief History of Computational Imaging	2
I-B	The Time-of-Flight Revolution	3
I-C	Roadmap of this paper	6
II	3D Imaging with ToF Sensors	6
II-A	Time-Domain 3D Imaging (TD-ToF)	7
II-B	Frequency-Domain 3D Imaging (FD-ToF)	9
III	Landscape of Inverse Problems in ToF Imaging	9
III-A	Optical Multi-path Interference and Spectral Estimation	10
III-B	Transient Imaging and Sampling Sparse Signals	11
III-C	Fluorescence Lifetime Imaging and System Identification	14
IV	Conclusions and Future Directions	17
	References	20
IV-A	Corresponding Author (Contact Details and Bio)	22
IV-B	Co-Author (Bio)	22

The authors are with the Massachusetts Institute of Technology, 75 Amherst Street, Cambridge, MA 02139 USA.

Abstract

Time-of-Flight (ToF) sensors offer a cost-effective and a real-time solution to the problem of three-dimensional imaging—a theme that has revolutionized our scene-understanding capabilities and is a topic of contemporary interest across many areas of science and engineering. The goal of this tutorial style article is to provide a thorough understanding of ToF imaging systems from a signal processing perspective that is useful to all application areas. Starting with a brief history of the ToF principle, we describe the mathematical basics of the ToF image formation process, for both time- and frequency-domain, present an overview of important results within the topic and discuss contemporary challenges where this emerging area can benefit from the signal processing community. In particular, we examine case-studies where inverse problems in ToF imaging are coupled with signal processing theory and methods, such as, sampling theory, system identification as well as spectral estimation, among others. Through this exposition, we hope to establish that ToF sensors are more than just depth sensors; depth information may be used to encode other forms of physical parameters, such as, the fluorescence lifetime of a bio-sample or the diffusion coefficient of turbid/scattering medium.

I. INTRODUCTION

A. A Brief History of Computational Imaging

“A picture is worth a thousand words.” Throughout the past several centuries, this immemorial phrase has pushed development of photography from analog to digital, accelerated by breakthrough advancements in sensor technology. The digital revolution created a necessity for sophisticated signal processing algorithms tailored for image enhancement, storage and compression. Shannon’s sampling theory was the pathway for analog-to-digital image conversion. The classical Wiener filter was used for image enhancement. Transform-domain coding utilizing discrete cosine and wavelet transforms played a pivotal role in JPEG compression. Over the last few decades, with the advent of the *wavelet transform* and *compressed sensing* theory, the field of signal processing has undergone a philosophical reformation. A field that was once heavily reliant on smoothness assumptions now uses principles inspired by the notion of sparsity. Until the last decade or so, the fields of image sensor technology and signal processing ran in parallel to one another with minimal interaction or exchange of know-how. However, there has recently been a growing trend towards the coherent co-design of sensors and algorithms: This is the theme of the emerging area of *Computational Imaging or Computational Photography* [1].

Practitioners of this computational imaging/photography ideology have devised many solutions which were previously not possible when individually adding functionality to the sensor architecture or using a more sophisticated processing algorithm. For example, consider the problem of motion deblurring that arises in conventional imaging. Whenever an object moves during sensor exposure, it causes pixels to smear across

the frame resulting in a blurred image. In the context of signal processing, this is an ill-posed problem that has been well-explored within the theme of deconvolution. The key problem is that the exposure time defines a temporal filter, which is essentially a box filter that annihilates any high-pass, spatial information. Consequently, algorithmic sophistication alone is not enough. Blurring may be avoided by a shorter exposure time, but this comes at the expense of low signal-to-noise ratio. It is clear that neither deconvolution nor the sensor level adaptation in itself suffice for a solution to the deblurring problem. The distinct role of computational imaging emerges when one considers the so-called *Flutter-Shutter* approach [2]. This approach involves a co-design of sensor and algorithms: in contrast to traditional imaging methods which require the shutter to be fully open during the exposure time, the flutter-shutter method flutters the shutter on and off in a binary, pseudo-random sequence. This sequence converts the non-invertible box-filter into an invertible one and, based on the choice of pseudo-random sequence, the corresponding deconvolution filter may be devised. Beyond deblurring in consumer imaging, the flutter-shutter approach is also useful in bio-imaging [3] where the imaging sensor may not be fast enough to capture flowing structures such as blood cells.

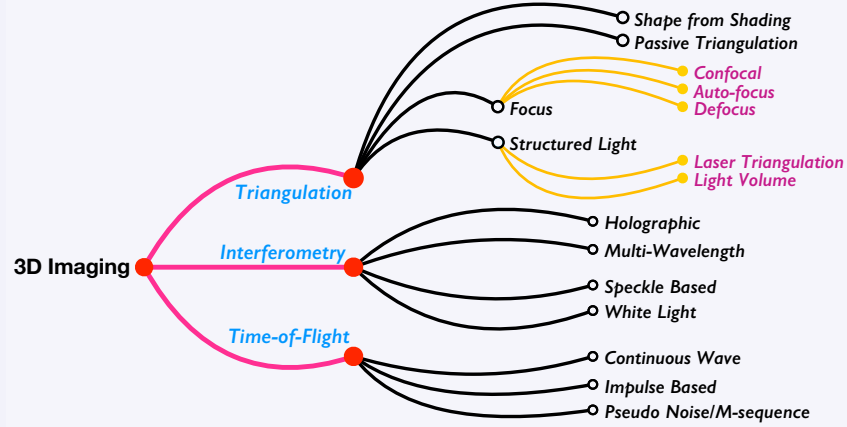
Other notable examples of the computational imaging philosophy are high dynamic range imaging [4], light-field imaging [5], [6], single pixel imaging [7] and Fourier ptychographic microscopy [8].

For the most part, image sensor design, signal/image processing and computational imaging have largely been restricted to two dimensional scenes. However, a true and richer representation of the environment around us lives in a three dimensional or 3D space. Capturing 3D information of a scene offers unparalleled benefits in accuracy and capabilities and is surely the future in many areas. This necessitates development of imaging modalities capable of recording 3D images.

A number of methods have been developed for the purpose of 3D imaging. An overview of the main techniques is presented in Box 1: Overview of 3D Imaging Techniques. Of all the 3D capture techniques, the Time-of-Flight (ToF) method has arguably attracted the most commercial and scientific interest in the last couple of years; there has been a surge of research towards improving both the sensor design as well as the algorithms used for processing 3D images. ToF imaging is the theme of this paper and we take a deep dive into the topic in the next section.

B. The Time-of-Flight Revolution

The ToF principle exploits the idea that distance and time are proportional quantities. As the name suggests, ToF is the round trip time between the source and the destination taken by a particle or a wave. Hence, knowing one entity is equivalent to knowing the other. Nature is replete with examples that rely on the ToF principle. Bats, dolphins and visually impaired human beings use the ToF principle for navigational purposes.

Box 1: Overview of 3D Imaging Techniques

Optical 3D scene capture can be broadly categorized into the following three approaches:

- ① *Triangulation* [9] method relies on the trigonometric principles. The distance to an unknown point is measured by computing the respective angles to the point from two edges of a triangle.
- ② *Time-of-Flight* or the ToF method, as the name suggests, relies on the time it takes for light to backscatter from an object at an unknown distance. Since time delays are linearly proportional to the distance, measuring the ToF amounts to measuring the range of the object.
- ③ *Interferometry* is similar to ToF with the main distinction that it requires the light waves to be coherent.

Each of these approaches can be further classified based on the application as well as the specialized principle linked with the approach.

Chronologically, the use of sound waves superseded the use of electromagnetic waves. Humans have used stones to estimate the depth of wells for millennia. The earliest work on using light waves for measuring ToF dates back to an experiment conducted by Galileo who was interested in estimating the speed of light. Unfortunately, his choice of distance (separation between two hills) did not lead to a conclusive result. The Danish astronomer Ole Rømer overcame this difficulty by using planetary distances. About two hundred years later, French physicist Hippolyte Fizeau was the first to estimate the speed of light precisely. Through the discovery of the photoelectric effect by Albert Einstein in the 1900s and the development of the electronic imaging sensors (CCD/CMOS), we are now at a point where the accumulated research efforts in the area of photonics and electronics have culminated in mass producible optical ToF sensors.

Contrary to conventional imaging sensors such as digital cameras that produce 2D images $I(x, y)$, ToF sensors capture 3D images, $I(x, y, z)$. The unique ToF sensor produces two images per exposure: an amplitude image and a depth image. The amplitude image is the standard 2D photograph, $I(x, y)$. Each pixel on the depth image represents the corresponding distance in the scene. The combination of the amplitude and the depth image produces the 3D image. We show the amplitude, depth and resulting 3D images in Fig. 1.

ToF based 3D imaging allows for applications that were previously unexplored. One of the first results

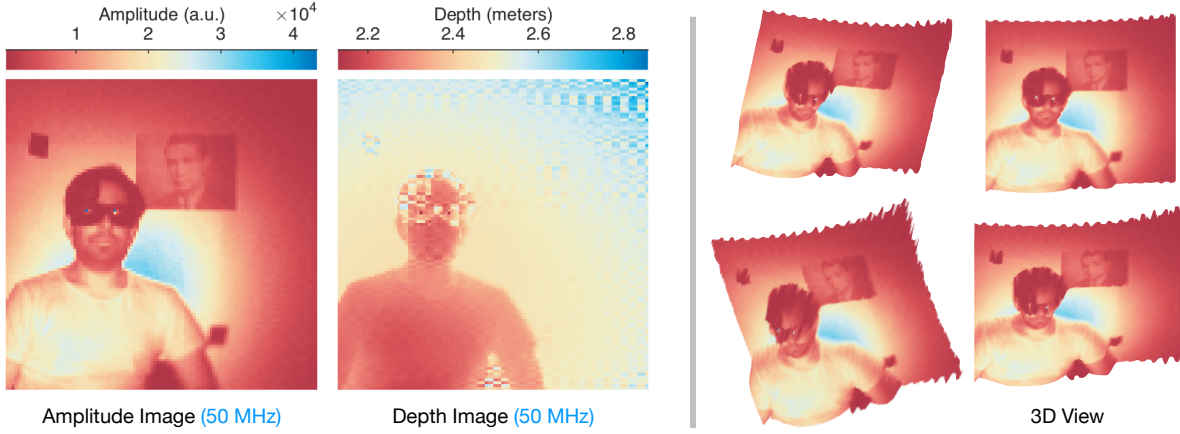


Fig. 1: 3D images captured via ToF sensor. We show amplitude image (or the conventional digital image), depth image and 3D images seen from multiple viewpoints.

demonstrated non-line-of-sight imaging capability [10]. This result—in parallel to “Doc” Edgerton’s famous *Bullet Through Apple*¹ image—lead to ultrafast imaging of light packets at an exorbitant frame rate of one trillion frames per second. A flurry of follow up work lead to results that allowed imaging through scattered media [11], light-in-flight imaging [12], [13] and 3D imaging in extremely low light [14].

With the advent of 3D sensing technology (most notably, the Microsoft Kinect XBox One) we can now replace room sized apparatus [10], moving sensors, and raster scan systems [14] by *miniaturized, cost-effective, real-time, and full-frame* ToF systems. Computational ToF imaging has already found a plethora of applications in, for example, ultra-fast imaging [15], [16], non line-of-sight imaging [17], imaging through scattering media, [18] and colored ToF imaging [19]. Outside of the computational imaging and human-computer interaction communities, an important application area is health care technology [20] and bio-imaging [21].

ToF sensors motivate a demand for forward models and algorithms that can handle this new wave of data. A handful of algorithms inspired by the signal processing community have been used to tackle inverse problems in optical ToF, but significant and numerous challenges still remain. Current efforts are directed at establishing empirical results with a rare discussion on design of *efficient algorithms, fundamental limits* or *performance bounds*.

While optical ToF sensors are a recent phenomenon, other ToF systems such as ultrasound, seismic and radar technology have been around for decades. The knowledge transfer between optical and other ToF systems is far from reality. Each ToF modality has its own idiosyncratic constraints which stem from the physics of the problem. However, there are commonalities that are shared by all of these systems. By discussing specific case studies, we take a first step towards the goal of bridging this gap.

¹<http://edgerton-digital-collections.org/galleries/iconic#hee-nc-64002>

C. Roadmap of this paper

The main goal of our tutorial style paper is to introduce our audience to the rapidly emerging field of ToF imaging from a signal processing perspective. We start this journey with a primer on the ToF image formation model. To reconcile any confusion stemming from taxonomical classification of ToF modalities—time domain (TD-ToF), frequency domain (FD-ToF) or amplitude modulated continuous wave (AMCW-ToF)—we discuss a unifying image formation model that is applicable to all known optical ToF systems. Furthermore, this model is backward compatible with other non-optical ToF modalities such as terahertz, ultrasound and microwave. Our first milestone covers a discussion on acquiring a single depth image: how consumer-grade ToF sensors capture 3D images? From there, the discussion turns to computational ToF imaging. Within the scope of computational ToF imaging, we enumerate case studies that map ToF imaging problems to signal processing subfields such as (a) line spectrum estimation theory, (b) sampling theory of sparse signals and (c) system identification. Finally, we discuss current challenges and future directions in computational ToF imaging where signal processing theory and methods can contribute significantly.

II. 3D IMAGING WITH ToF SENSORS

ToF sensors are active illumination devices consisting of an illumination unit capable of probing a scene with an amplitude modulated light that is not necessarily coherent. We call this amplitude modulated waveform the *probing function* or $p(x, y, t)$ where (x, y) are the spatial co-ordinates and t is the continuous time variable. For simplicity of exposition, we will consider per-pixel processing and simply write $p(t)$. The probing function interacts with the *scene response function* (SRF). This interaction results in the *reflected signal* $r(t)$. The reflected signal is observed at the ToF sensor, which is characterized by its transfer function that we refer to as the *instrument response function* (IRF). The IRF models the sensor's electro-optical assembly. For example, the IRF for a digital camera is the point spread function of the lens. The interaction between the reflected signal and the ToF sensor results in the measured signal $m(t)$ which is converted to a digital signal via sampling. Precise mathematical description of the ToF imaging pipeline is discussed in Box 2: ToF Image Formation Model. Based on the ToF image formation model parameters, $\{p, h, \varphi\}$, one may now define the specific inverse problem at hand. This is also true of other wave-based ToF modalities—radar, sonar, ultrasound, terahertz, and others. A distinct feature of the consumer ToF sensors is their use of the lock-in principle [22] which implements the cross-correlation operation. From a mathematical standpoint, and in the absence of noise and distortion, this translates to the fact that,

$$\varphi(t, \tau) = p(\tau + t). \quad (1)$$

Next, we discuss how consumer ToF sensors capture 3D images. Depending on the choice of probing function, the ToF imaging setup may be categorized into time-domain or frequency-domain modes. In either case, the SRF for the case of single depth imaging is modeled as a shift-invariant kernel,

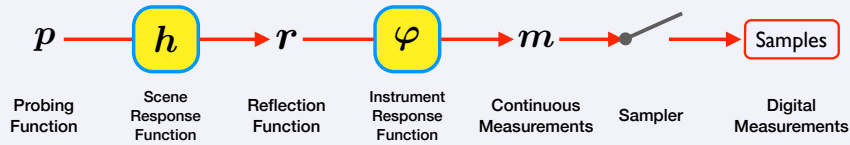
$$h(t, \tau) = \Gamma_0 \delta(t - \tau - t_0), \quad t_0 = 2 \frac{d_0}{c} \quad (4)$$

where the goal is to estimate the amplitude and delay (or depth), $\{\Gamma_0, t_0\}$, respectively, at each pixel.

A. Time-Domain 3D Imaging (TD-ToF)

TD-ToF systems probe the scene with a time-localized, periodic signal of form $p(t) = p(t + \Delta)$, $\Delta > 0$. While specialized scientific instruments, such as the streak tube, may be able to produce a pulse that mimics the Dirac's Delta distribution δ [10], this form of precision is impractical for consumer-grade instruments. In practice, a maximum-length sequence (MLS) [23] is an optimal choice of probing function in regards to

Box 2: ToF Image Formation Model



Elements of the ToF imaging pipeline:

- 1) Probing Function denoted by $p(t)$ represents the waveform emitted by the ToF sensor's illumination unit. The probing function may be a time-localized pulse or a continuous wave. In either case, it is chosen to be a periodic function of form $p(t) = p(t + \Delta)$, $\Delta > 0$.
- 2) Scene Response Function or SRF denoted by $h(t, \tau)$ models the transfer function of the scene. For example, for an object with reflection coefficient Γ and at a depth d meters away from the sensor, the SRF takes form of, $h(t, \tau) = \Gamma \delta(t - \tau - 2d/c)$. Here c is the speed of light. The SRF may also be characterized as the Green's function of a differential equation that models some physical phenomenon such as fluorescence emission, diffusion or scattering.
- 3) Reflected Function denoted by $r(t)$ is the result of interaction between the probing signal and the SRF. The reflected signal is modeled as a Volterra/Fredholm integral,

$$r(t) = \int_{\Omega_1} p(\tau) h(t, \tau) d\tau. \quad (2)$$

Whenever the SRF is a shift-invariant kernel, i.e., $h_{SI}(t, \tau) = h(t - \tau)$, the reflected signal is simply a convolution/filtering operation between the probing function and the SRF, $r(t) = (p * h)(t)$.

- 4) Instrument Response Function or the IRF denoted by $\varphi(t, \tau)$ models the transfer function of the electro-optical elements of the ToF sensor. For example, in conventional digital cameras, the spatial IRF is the point spread function of the lens.
- 5) Measurements denoted by $m(t)$ are a result of sensing the reflected signal via the electro-optical elements of the ToF sensor. Continuous-time measurements are modeled as,

$$m(t) = \int_{\Omega_2} r(\tau) \varphi(t, \tau) d\tau. \quad (3)$$

The ToF sensor stores discrete measurements by sampling continuous-time signal $m(t)$ and this results in the discrete sequence $m[k] = m(kT)|_{t=kT, k \in \mathbb{Z}}$ where $T > 0$ is the sampling interval.

In many practical cases of interest, both the SRF and the IRF are shift-invariant. In that case, the measurements can be written as a convolution product, $m(t) = (p * h * \varphi)(t)$. Whenever the IRF is a function of form $\varphi(t, \tau) = \varphi(t + \tau)$, the measurements amount to $m(t) = (r \otimes \varphi)(t)$ where $\otimes$ denotes cross-correlation operation. "Lock-in" sensors operate on this principle.

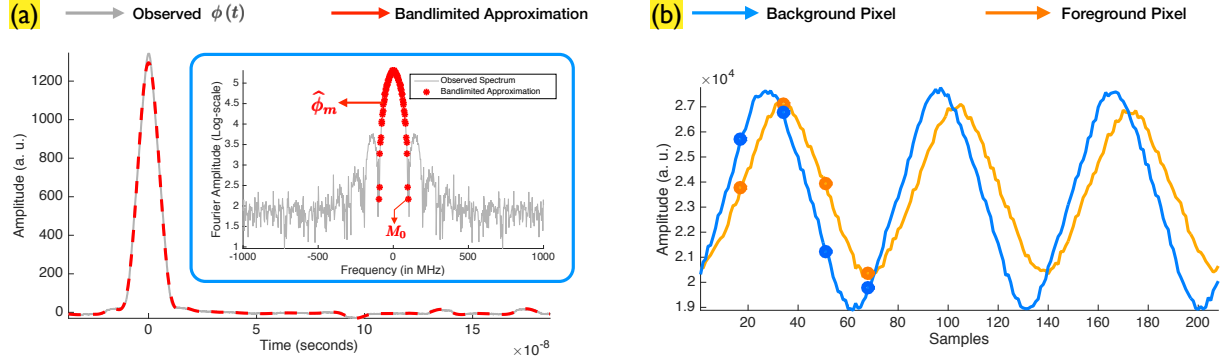


Fig. 2: Time and Frequency Domain ToF signals. (a) Bandlimited approximation of auto-correlated probing signal ($\phi = p * \bar{p}$) in time-domain ToF setup. The low-pass property is evident from its Fourier spectrum. This is a result of experiment with $\Delta = 310$ ns and $M_0 = 30$. (b) Samples of measured signal in (6) which were used to create the depth map in Fig. 1. We plot 208 samples for an experiment with $f = \omega/2\pi = 50$ MHz. With reference to Fig. 1, the foreground and background pixels map to 2.2 and 2.5 meters, respectively.

time-localization. In this case, given (2) and (4), the reflected signal reads $r(t) = \Gamma_0 p(t - t_0)$, with delay $t_0 = 2d_0/c$. Due to the lock-in sensor architecture [22], which constraints the IRF in (1), the measurements simplify to $m(t) = (r \otimes p)(t) \equiv (\bar{r} * p)(t)$ where $\otimes$ denotes the cross-correlation operation and $\bar{r}(t) = r(-t)$. A closer look at m reveals an underlying auto-correlation operation involving p , that is, $\bar{m}(t) = \Gamma_0 (p * \bar{p})(t - t_0)$. Consequently, we may write $\phi = p * \bar{p}$. The ToF is then estimated by solving for $\tilde{t}_0 = \arg \max_{t_0} \bar{m}(t) = \arg \max_{t_0} \phi(t - t_0)$.

Whenever $p(t)$ is modeled to be some parametric waveform, such as a Gaussian function, B-spline or a combination of parametric pulses, parameter estimation techniques may be used to estimate the ToF t_0 , and the reflection coefficient Γ_0 . However, this may not be the case in practice because of model mismatch or the physical aspects of light propagation. In such a setting, it is effective to use the property of *bandlimited approximation*: approximate p , and hence ϕ , with, the first few Fourier components,

$$\tilde{p}(t) = \sum_{|m| \leq M_0} \hat{p}_m e^{jm\omega_0 t} \quad \text{with} \quad \hat{p}_m = \frac{1}{\Delta} \int_0^\Delta p(t) e^{-jm\omega_0 t} dt, \quad (5)$$

where $\omega_0 = 2\pi/\Delta$ is the fundamental frequency and Δ is the maximum operating range of the ToF sensor. This choice is aptly justified if one considers the fact that,

- ♦ Most electronic/optical instrumentation is approximately bandlimited due to physical constraints [24].
- ♦ The probing function may not admit a parametric representation. Even if the probing function assumes a parametric representation, bandlimited approximation via Fourier series coefficients circumvents the estimation of parameters of the probing function.

The utility of bandlimited approximation property is demonstrated via experiments shown in Fig. 2(a). Starting with an MLS, we design a probing function. We plot $\phi = p * \bar{p}$ together with its bandlimited approximation

$\tilde{\phi}(t)$ obtained by retaining first M_0 Fourier series coefficients $\left\{\hat{\phi}_m = |\hat{p}_m|^2\right\}_{|m| \leq M_0}$. We are thus able to re-write measurements as $\bar{m}(t) = \Gamma \phi(t - t_0) \equiv C_0 \sum_{|m| \leq M_0} \hat{\phi}_m e^{jm\omega_0 t}$ where the complex-valued constant $C_0 = \Gamma e^{-j\omega_0 t_0}$ is the unknown to be estimated.

B. Frequency-Domain 3D Imaging (FD-ToF)

ToF sensors such as the Microsoft Kinect XBox One use a continuous wave based probing function $p(t) = 1 + p_0 \cos(\omega t)$, $p_0 < 1$ where ω is the modulation frequency and p_0 is the modulation amplitude. With the SRF defined in (4), the reflected signal reads $r(t) = \Gamma p(t - t_0)$, $t_0 = 2d_0/c$. Again, the lock-in sensor acts as an electronic homodyne detector such that (1) holds, and measurements results in [25],

$$m(t) \stackrel{(3)}{=} \Gamma_0 \left(1 + \frac{p_0^2}{2} \cos(\omega t + \theta_\omega)\right), \quad \theta_\omega \in [0, 2\pi) \quad (6)$$

where $\theta_\omega = \omega t_0 = 2d_0\omega/c$ is the frequency dependent phase. The ToF sensor records discrete measurements of form, $m_k = m(kT)$ with $T = \pi/2\omega$ and uses a phase estimation algorithm commonly known as the ‘‘Four Bucket Method’’ [21], [22] to estimate parameters d_0 and Γ_0 . For a given modulation frequency, this method works with 4 discrete measurements $\{m_k\}_{k=0}^3$ that are used to form a complex number $z_\omega \in \mathbb{C}$,

$$z_\omega = (m_0 - m_2) + j(m_3 - m_1) \text{ with } \begin{bmatrix} m_0 & m_1 \\ m_2 & m_3 \end{bmatrix} = \frac{\Gamma_0}{2} \begin{bmatrix} 2 + p_0^2 \cos(\omega t_0) & 2 - p_0^2 \sin(\omega t_0) \\ 2 - p_0^2 \cos(\omega t_0) & 2 + p_0^2 \sin(\omega t_0) \end{bmatrix}. \quad (7)$$

The scene parameters are then estimated by $\tilde{\Gamma}_0 = |z_\omega|/p_0^2$ and $\tilde{d}_0 = c\angle z_\omega/2\omega$. We use this method to create the depth image in Fig. 1. The raw data-samples corresponding to the experiment are plotted in Fig. 2(b). This completes our discussion on depth imaging with time- and frequency-domain ToF sensors.

Key takeaways of this section:

- ▶ ToF imaging pipeline consists of a probing function, scene response function and an instrument response function.
- ▶ For the case of 3D imaging, the scene response function is the shift-invariant, time delay operator in (4).
- ▶ In time-domain ToF (TD-ToF), the probing function is a time-localized pulse (cf. 2(a)) and the object’s distance from the sensor is encoded as the round-trip time-delay $t_0 = 2d_0/c$.
- ▶ In frequency-domain ToF (FD-ToF), the probing function is a sinusoidal waveform with known modulation frequency. The object’s distance from the sensor is encoded as a frequency dependent phase.

III. LANDSCAPE OF INVERSE PROBLEMS IN TOF IMAGING

Having covered the mathematical basics of time- and frequency-domain ToF imaging systems, we will now discuss case studies where signal processing theory may be used in combination with ToF sensors to solve inverse problems.

A. Optical Multi-path Interference and Spectral Estimation

Multi-path interference or MPI is a problem that naturally occurs in communications, acoustics and array signal processing. Consumer-grade FD-ToF sensors are designed to work under the premise that each point in the scene maps to a single pixel on the sensor. Whenever this hypothesis does not hold, the ToF sensor measurements are erroneous due to MPI. This results in corrupted 3D images. For example, this can happen when imaging through a semi-transparent object such as a glass window.

A consequence of multiple optical paths combining at a given pixel is that the shift-invariant SRF for single depth (4) now takes the form of a K -sparse filter,

$$h_K(t) = \sum_{k=0}^{K-1} \Gamma_k \delta(t - t_k), \quad t_k = 2 \frac{d_k}{c}, \quad (8)$$

where K is the number of optical paths, and $\{\Gamma_k, d_k\}_{k=0}^{K-1}$ are the scene parameters corresponding to each optical path. For simplicity, we assume that $p_0 = 1$ in the definition of the probing function, $p(t) = 1 + p_0 \cos(\omega t)$. The reflected signal in this case reads,

$$r_K(t) = \sum_{k=0}^{K-1} \Gamma_k \cos(\omega t - \omega t_k) = \sum_{k=0}^{K-1} \Gamma_k \cos(\omega t - \theta_{k,\omega}), \quad (9)$$

which indicates an addition of sinusoids with varying phases. Let us resort to complex-valued representation and let z^* denote the complex-conjugate of z . The measurements take the form,

$$m_K(t) = \frac{1}{2} e^{j\omega t} \underbrace{\sum_{k=0}^{K-1} \Gamma_k e^{+j\omega t_k}}_{\hat{h}_K^*(\omega)} = \frac{1}{2} e^{j\omega t} \hat{h}_K^*(\omega) \quad \text{with} \quad \underbrace{\hat{h}_K(\omega) = \sum_{k=0}^{K-1} \Gamma_k e^{-j\omega t_k}}_{\text{Fourier Transform}} \quad (10)$$

where $\hat{h}_K(\omega)$ is the Fourier Transform of the shift-invariant SRF in (8)—a sum of K complex exponentials or phasors. Since the probing function is a sinusoid—an eigen-function of a linear system—the measurements amount to observing Fourier Transform at modulation frequency ω .

Whenever $K = 1$ —the case of single depth 3D imaging with no MPI—the unknown constant $\hat{h}_1(\omega) = \Gamma_0 e^{-j\omega t_0}$ is directly estimated by implementing the four bucket method that maps $\{m_K(\pi k/2\omega)\}_{k=0}^3 \rightarrow z_\omega = \hat{h}_1^*(\omega)$. However, in presence of MPI, the FD-ToF sensor estimates the scalar $z_\omega = \hat{h}_K^*(\omega)$ rather than $\{\Gamma_k, d_k\}_{k=0}^{K-1}$ —the pixel-wise scene parameters. For $K = 2$, we plot the SRF in time-domain, frequency-domain and phasor-domain in Fig. 3(a). The fact that $z_\omega = \hat{h}_K^*(\omega)$ is an argument of modulation frequency motivates the necessity for frequency diversity [25]. Given N harmonic measurements of form $\{z_{n\omega_0}\}_{n=0}^{N-1}$, the task of estimating $2K$ unknowns $\{\Gamma_k, d_k\}_{k=0}^{K-1}$ is a classical problem in signal processing which is studied under the theme of *Spectral Estimation Theory* [26]. Due to the pervasiveness and wide applicability of this problem, a self-contained review of Prony's method that seeks to estimate $\{\Gamma_k, d_k\}_{k=0}^{K-1}$ given $\{z_{n\omega_0}\}_{n=0}^{N-1}$ is

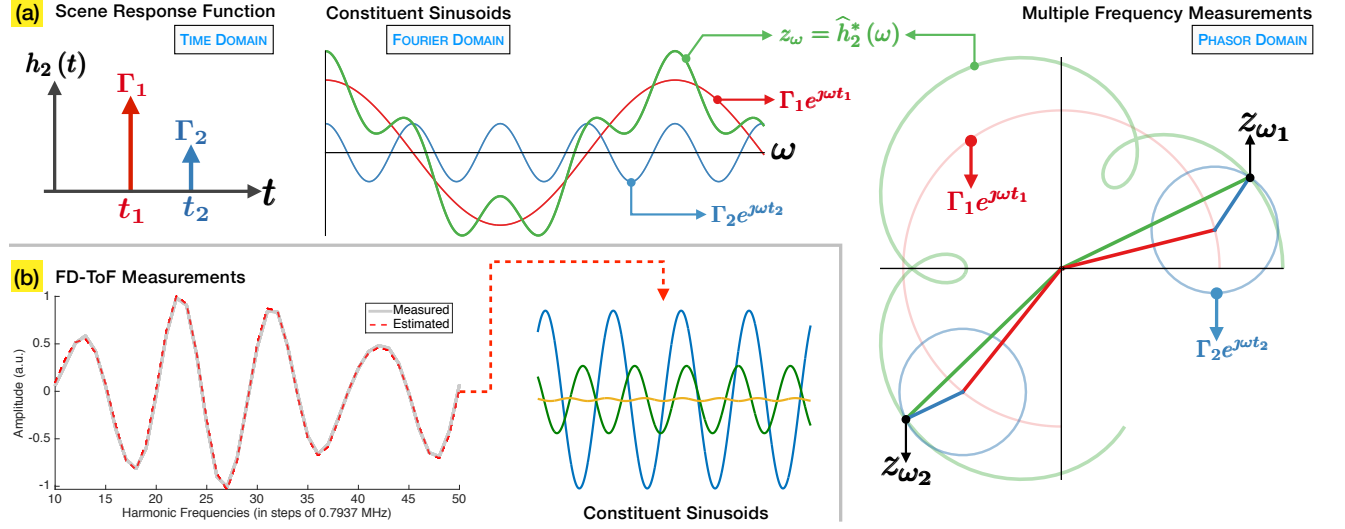


Fig. 3: Multi-path interference in ToF sensors. (a) Scene Response Function for $K = 2$ in time-domain (cf. (8)), frequency-domain and complex phasor-domain. At a given modulation frequency ω , the ToF sensor measures the complex number z_ω . However, the correct values to be estimated are $\{\Gamma_k, d_k\}_{k=0}^{K-1}$. This model mismatch results in erroneous depth reconstruction. (b) Experimentally measured multi-frequency data and its decomposition into sinusoids via spectral estimation.

provided in Box 3: Spectral Estimation. This approach may be used for correction of MPI in ToF sensors. In Fig. 3(b), we show multi-frequency data acquired using a ToF sensor with $K = 3$. Our demonstration that is based on the matrix pencil method [26] shows the constituent sinusoidal components. Other methods such as the orthogonal matching pursuit [25] or EPIRIT and MUSIC [26] may also be used.

While MPI suggests that interfering optical paths are a nuisance, it may be exploited for 3D scanning and imaging of translucent objects. We demonstrate *multiple depth imaging* capability of ToF sensors in Fig. 4 where upto $K = 3$ optical paths interfere at the sensor. As seen in Fig. 4, the resulting 3D images for various modulation frequencies are erroneous. Although Prony's method for spectrum estimation is highly unstable in presence of noise and model mismatch, oversampling—that is, using much more than $2K$ measurements—is a reasonable solution to counteract perturbations in data. Again, we use the matrix pencil method to estimate the multiple depth related scene parameters. With 46 multiple frequency measurements (cf. Fig. 3(b)), the estimation procedure shows that it is possible to reconstruct objects at multiple depths even in presence of MPI. The recovered amplitude and phase images, $\{\Gamma_k\}_{k=0}^2$ and $\{t_k\}_{k=0}^2$, respectively, are shown in Fig. 4.

B. Transient Imaging and Sampling Sparse Signals

Understanding light propagation through physical medium has interesting theoretical and practical consequences. For example, a recent demonstration by Velten and co-workers [10] showed that the information contained in multiple echoes of light can be used for non-line-of-sight imaging. From a signal processing perspective, the properties of light/wave propagation can be best understood as a time-dependent transfer

Box 3: Spectral Estimation (Prony's Method)

Given $2K$ contiguous measurements $z_{n\omega_0} = \sum_{k=0}^{K-1} \Gamma_k e^{jn\omega_0 t_k}$, $n \in [m, 2K+m)$, Prony's Method aims to estimate unknowns $\{\Gamma_k, d_k\}_{k=0}^{K-1}$ from a system of equations that is linear in reflection coefficients Γ_k but non-linear in depths $t_k = 2d_k/c$. The solution to this problem relies on an underlying *annihilation equation*. Let us define a polynomial of degree K ,

$$Q(z) = \prod_{k=1}^K (z - e^{j\omega_0 t_k}) = z^K \sum_{m=0}^K q_m z^{-m}$$

whose roots are $\{e^{j\omega_0 t_k}\}_{k=0}^{K-1}$. Observe that filtering sequence $\{q_m\}_{m=0}^K$ with measurements $\{z_{m\omega_0}\}_{m=0}^{K-1}$ results in nullity,

$$(q_m * z_{m\omega_0}) = \sum_{\ell=0}^K q_\ell z_{\omega_0(m-\ell)} = \sum_{\ell=0}^K q_\ell \sum_{k=0}^{K-1} \Gamma_k e^{j(m-\ell)\omega_0 t_k} = \sum_{k=0}^{K-1} \Gamma_k \underbrace{\left(\sum_{\ell=0}^K q_\ell e^{-j\ell\omega_0 t_k} \right)}_{Q(e^{j\omega_0 t_k})=0} e^{jm\omega_0 t_k} = 0$$

Given $\{z_{n\omega_0}\}_{m \leq n < m+2K}$, in vector-matrix notation, solution to the above equation is achieved by solving for,

$$m \leq n < m+2K, \quad \underbrace{\begin{bmatrix} z_{(m+K)\omega_0} & z_{(m+K-1)\omega_0} & \cdots & z_{m\omega_0} \\ z_{(m+K+1)\omega_0} & z_{(m+K)\omega_0} & \cdots & z_{(m+1)\omega_0} \\ \vdots & \vdots & \ddots & \vdots \\ z_{(m+2K-1)\omega_0} & z_{(m+2K-2)\omega_0} & \cdots & z_{(m+K-1)\omega_0} \end{bmatrix}}_{\text{Toeplitz Matrix}} \begin{bmatrix} q_0 \\ q_1 \\ \vdots \\ q_K \end{bmatrix} = 0.$$

With filter $\{q_m\}_{m=0}^{m=K}$ known, we construct $Q(z)$ and its roots lead to the K estimates, $\{t_k\}_{k=0}^{K-1}$. With the depths known, estimating $\{\Gamma_k\}_{k=0}^{K-1}$ is the usual linear least-squares problem since Γ_k 's are linearly constrained in definition of $z_{n\omega_0}$.

function of the scene. The three main ingredients of time-resolved light transport are direct reflections, inter-reflections and sub-surface scattering:

$$\underbrace{\Gamma_0 \delta(t - t_0)}_{\text{Direct Reflection}} + \underbrace{\sum_{k=1}^K \Gamma_k \delta(t - t_k)}_{\text{Inter-Reflections}} + \underbrace{\Gamma_s \delta(t - t_s) * e^{-\mu t} u(t)}_{\text{Sub-surface Scattering}}$$

where $\{\Gamma_k, t_k\}_{k=0}^{K-1}$ and $\{\Gamma_s, t_s\}$ are the usual multi-path components, μ is the absorption coefficient and $u(t)$ is the Heaviside function. If it were possible to create a highly time-localized probing function of form $p \sim \delta$, identification of scene related parameters, $\{\{\Gamma_k, t_k\}_{k=0}^{K-1}, \Gamma_s, t_s, \mu\}$ would be a relatively easy task. However, the time-resolution of probing functions programmed on the state-of-the-art ToF sensors is, at best, about 11 nanoseconds (cf. Fig. 2(a)). This is orders of magnitude longer than the sophisticated apparatus used by Velten et al. [10] which attains a pulse width of approximately 50 picoseconds. That said, TD-ToF sensors may still be used for imaging transient phenomenon linked with direct and indirect inter-reflections. In fact, this problem is closely tied with the problem of sampling of continuous-time sparse signals [27].

The SRF for the transient imaging problem with direct and indirect components of the light propagation is a continuous-time K sparse signal,

$$h_K(t) \stackrel{(8)}{=} \underbrace{\Gamma_0 \delta(t - t_0)}_{\text{Direct Reflection}} + \underbrace{\sum_{k=1}^K \Gamma_k \delta(t - t_k)}_{\text{Indirect Reflections}}.$$

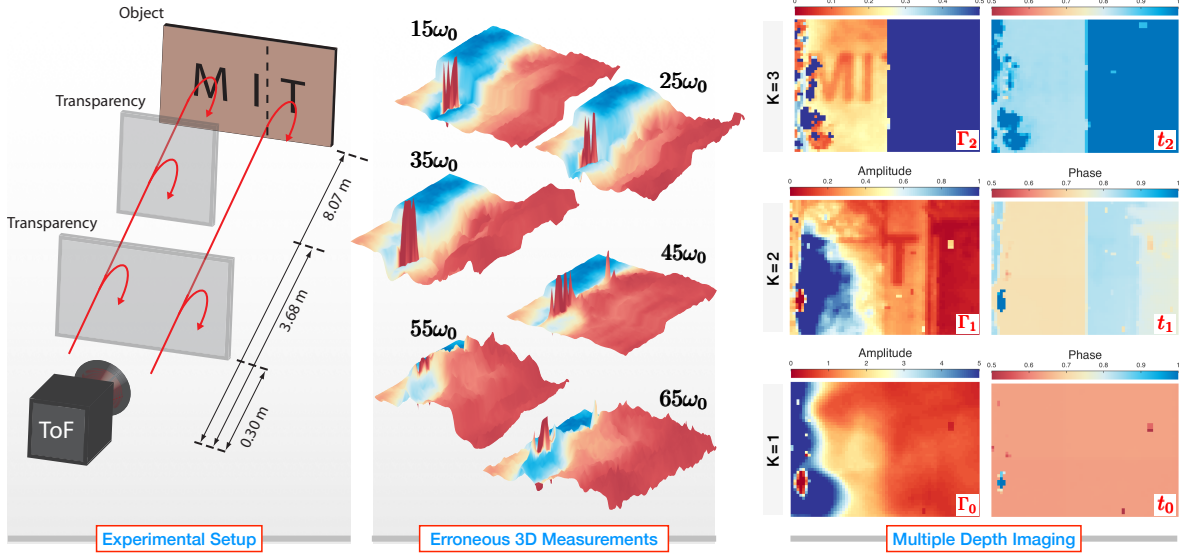


Fig. 4: Multi-depth Imaging with ToF sensors. The experimental setup describes a scene with upto $K = 3$ optical paths arising from 3 different depths. With $f_0 = \omega_0/2\pi = 0.8$ MHz, raw 3D images using the four bucket method. The resulting 3D maps are erroneous. We show 3D images for modulation frequencies $15\omega_0$ to $65\omega_0$ in steps of $10\omega_0$. We estimate $\{\Gamma_k, t_k\}_{k=0}^2$ using spectrum estimation method implemented via matrix pencil algorithm. The results are consistent with the experimental setup. For example, the letters “MI” and “T” associated with Γ_1 and Γ_2 correspond to their respective depths.

With the IRF defined in (1), the ToF sensor measurements are $m(t) = (r \otimes p)(t) = (p * \bar{p} * \bar{h}_K)(t)$ (cf. Box 2). With the commutativity property of the convolutions, the sampled TD-ToF measurements may be written as $m_n = (\phi * \bar{h}_K)(t)|_{t=nT}, n = 0, \dots, N-1$ where $\phi(t) = (p * \bar{p})(t)$. Since the probing function admits a bandlimited approximation (see (5) and Fig. 2(a)), the time-reversed TD-ToF measurements may be reinterpreted as samples of a sparse signal that has been low-pass filtered with a calibrated sampling kernel $\phi(t)$. Furthermore, the bandlimited approximation property allows us to write,

$$\bar{m}_n = (\phi * h_K)(t)|_{t=nT} \approx \sum_{\ell=-M_0}^{\ell=+M_0} \hat{\phi}_\ell \hat{h}_K(\ell\omega_0) e^{j\ell\omega_0 nT}, \quad n = 0, \dots, N-1,$$

where $\hat{\phi}_\ell = |\hat{p}|^2$ since the sampling kernel ϕ is obtained by auto-correlating the probing function, and $\hat{h}_K(\ell\omega_0)$ is the sampled Fourier Transform of the SRF. Note that the Discrete Fourier Transform of the sampled measurements $\{\bar{m}_n\}_{n=0}^{N-1}$ results in the $2M_0+1$ point-wise samples of the weighted Fourier Transform of the SRF, $\{\hat{\phi}_\ell \hat{h}_K(\ell\omega_0)\}_{|\ell| \leq M_0}$. Hence, given sampled measurements we can deconvolve the SRF in Fourier domain, which is a sum of K complex exponential functions. The only requirement being that $\hat{\phi}_\ell$ should not vanish in the interval $-M_0 \leq \ell \leq M_0$. In a noiseless setting, upon deconvolving the SRF $\{\hat{h}_K(\ell\omega_0)\}_{|\ell| \leq M_0}$, we can use spectral estimation (cf. Box 3) to estimate the parameters $\{\Gamma_k, t_k\}_{k=0}^{K-1}$, provided that $M_0 \geq K$. Several approaches have been proposed in literature to stabilize the spectral estimation method in presence of noise [26]. As before, we use the matrix pencil method for our experimental demonstrations.

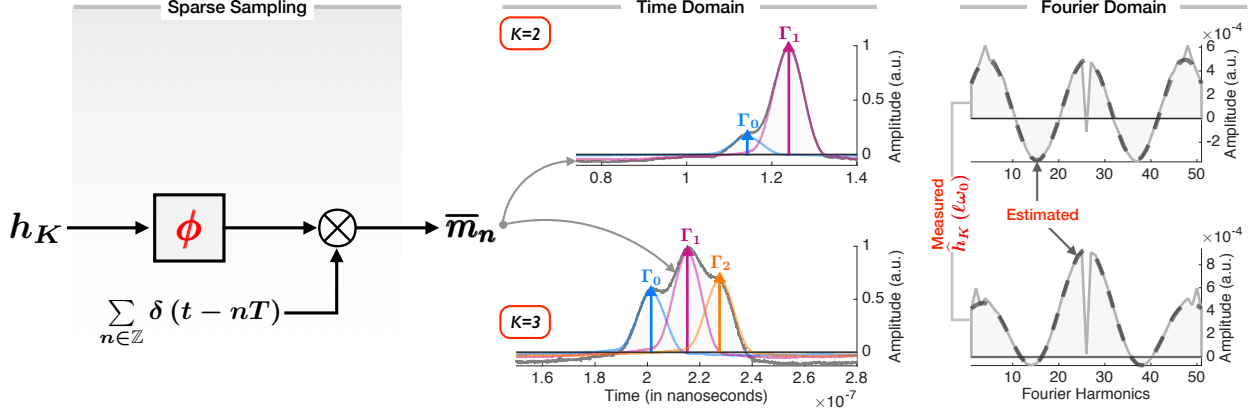


Fig. 5: Time-domain ToF can be reinterpreted as a sparse sampling problem in case of transient imaging. The goal is to recover the sparse scene response function given its low-pass filtered samples with bandwidth $M_0 = 25$. In context of sampling theory, the low-pass kernel ϕ is an auto-correlation of the probing function. We discuss recovery of sparse signal for the case $K = 2$ and $K = 3$. We also show the real part of the measured and estimated Fourier Transform of the scene response function $\hat{h}_K(\ell\omega_0)$.

Fig. 5 illustrates the reinterpretation of the transient imaging problem as a sparse sampling problem. Time-domain measurements are acquired for the case $K = 2$ and $K = 3$, respectively, with sampling rate $T = 70$ picoseconds. In each case, we recover the K -sparse SRF, h_K in following steps,

- 1) Obtain the Discrete Fourier Transform of sampled measurements $\bar{m}_n$, that is $\hat{m}_\ell^*$.
- 2) Deconvolve $\hat{m}_\ell^*$ with the sampling kernel $\hat{\phi}_\ell$ to obtain, $\hat{h}_K(\ell\omega_0) = \hat{m}_\ell^* / \hat{\phi}_\ell$, $\ell \in [-M_0, M_0]$, $M_0 \geq K$.
- 3) Use Prony's method (cf. Box 3) to estimate the sparse SRF $\{\Gamma_k, t_k\}_{k=0}^{K-1}$ given $\hat{h}_K(\ell\omega_0)$.

We plot the sampled, time-reversed measurements $\bar{m}_n$ together with the real part of $\hat{h}_K(\ell\omega_0)$ in Fig. 5. The resulting SRF parameters $\{\Gamma_k, t_k\}_{k=0}^{K-1}$ are consistent with the experimental setup.

Imaging transient phenomenon may be used for several applications. For example, consider the experimental setup in Fig. 6. A diffusive, semi-translucent sheet covers a placard that reads “Time of Flight.” A conventional digital camera is not be able to image through the diffuse/scattering object. However, in the context of a TD-ToF imaging setup, this scene has a sparse SRF with $K = 2$. In fact, the case of $K = 2$ described in Fig. 5 uses measurements directly from the example in Fig. 6, under consideration. By applying the sparse signal recovery method discussed in this section, it is possible to estimate the SRF h_2 . We discuss the results of this experiment in Fig. 6. As can be seen, the amplitude image corresponding to Γ_1 matches the experimental setup, and we are thus able to image through a diffusive medium.

C. Fluorescence Lifetime Imaging and System Identification

Fluorescence lifetime imaging (FLI) is a problem of interest across several disciplines of science and engineering, with the notable example of fluorescence lifetime microscopy. This imaging modality finds

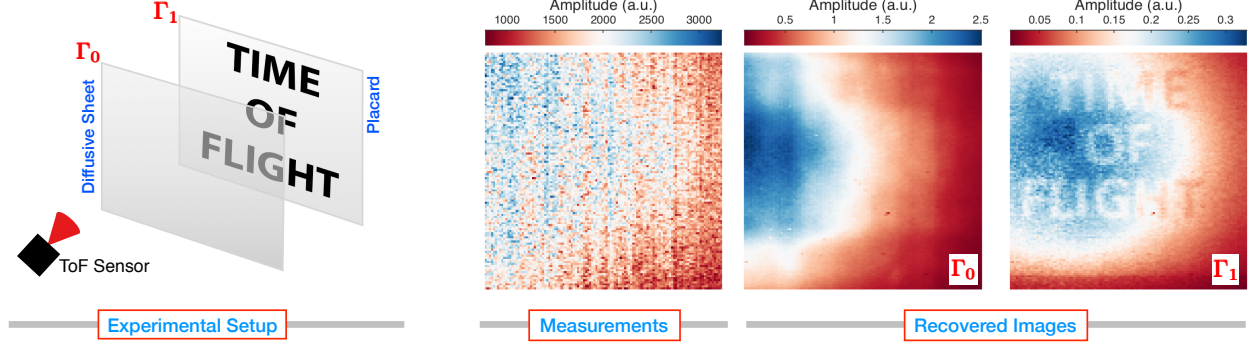


Fig. 6: Experimental setup for $K = 2$, TD-ToF measurements and transient image components Γ_0 and Γ_1 .

applications in a number of fundamental problems including DNA sequencing, tumor detection, fluorescence tomography and high resolution microscopy. Like the ToF sensors, FLI can be categorized into time-domain and frequency-domain modes. Time-domain FLI (TD-FLI) utilizes an impulse-like excitation pulse which is used to probe a fluorescent sample. The resulting time-resolved, reflected signal is then used to calculate lifetimes. Given a fluorescent sample with lifetime parameter λ_0 , this interaction is modeled as, $m(t) = \delta(t) * \rho_0 e^{-\frac{t}{\lambda_0}} u(t)$, where m is the measured signal, δ is the Dirac distribution and u is the usual Heaviside function. On the other hand, in case of frequency-domain FLI (FD-FLI), the sample is excited with a sinusoidal probing function of form $\cos(\omega t)$. The phase of the reflected signal ($\propto \cos(\omega t - \tan^{-1}(\omega \lambda_0))$) encodes the lifetime parameter λ_0 . Let $h(t)$ be a linear, shift-invariant system. Recall that,

$$r(t) = \cos(\omega t) * h(t) = |\hat{h}(\omega)| \cos(\omega t + \angle \hat{h}(\omega)), \quad (11)$$

where $\hat{h}$ is the Fourier Transform of h . More precisely, with $h(t) = \rho_0 e^{-\frac{t}{\lambda_0}} u(t)$, its Fourier Transform is $\hat{h}(\omega) = \frac{\rho_0 \lambda_0}{1 + j\omega \lambda_0}$, then $\angle \hat{h}(\omega) = -\tan^{-1}(\omega \lambda_0)$. Consequently, the FD-FLI problem boils down to phase estimation. This is a non-linear problem and several methods have been proposed in literature. Noting that the measurements (11) are proportional to the so called *modulation depth*, that is, $|\hat{h}(\omega)|$, the intensity of the reflected signal may also be used for lifetime estimation due to its dependence on λ_0 .

In either case, time- or frequency-domain FLI, the equipment is prohibitively expensive due to strict system constraints and precise electro-optical components. Furthermore, measurements must be calibrated in order to account for path delays attributed to sample's placement relative to the imaging sensor. Alternatively, ToF sensors are a consumer-grade commodity and are available at an affordable price. In context of ToF imaging, the SRF for FLI is a shift-invariant function that takes form of [21],

$$h_{\text{FLI}}(t) = \underbrace{\Gamma_0 \delta(t - t_0)}_{\text{Direct Reflection}} + \underbrace{\rho_0 e^{-\frac{t-t_0}{\lambda_0}} u(t - t_0)}_{\text{Fluorescent Sample}}, \quad t_0 = 2\frac{d_0}{c}, \quad (12)$$

where d_0 is the distance of the sample from the sensor—a quantity that is often calibrated in FLI setup,

which we assume to be an unknown. In parallel to the frequency-domain FLI, the probing function in case of FD-ToF is defined by $p(t) = 1 + p_0 \cos(\omega t)$. With the IRF defined in (1), the measurements read

$$m(t) \stackrel{(6)}{=} \left| \hat{h}_{\text{FLI}}(0) \right| + \left| \hat{h}_{\text{FLI}}(\omega) \right| \frac{p_0^2}{2} \cos\left(\omega t - \angle \hat{h}_{\text{FLI}}(\omega)\right), \quad (13)$$

where $\hat{h}_{\text{FLI}}(\omega)$ is the Fourier Transform of the SRF. As described in Section II-B, the ToF sensor uses the four bucket method (7) to record the sampled SRF spectrum at modulation frequency ω ,

$$z_\omega = p_0^2 \hat{h}_{\text{FLI}}^*(\omega) \quad \text{where} \quad \hat{h}_{\text{FLI}}(\omega) = \left(\Gamma_0 + \frac{\rho_0 \lambda_0}{1 + j\omega \lambda_0} \right) e^{-j\omega t_0}. \quad (14)$$

We show the amplitude and phase images, $|z_\omega|$ and $\angle z_\omega$, respectively in Fig. 7(a) with $f = \omega/2\pi = 20$ MHz. As shown in [21], the reflection coefficients are dependent on the optical frequency of light and the direct component may be filtered ($\Gamma_0 = 0$). Since the ToF sensor samples the Fourier spectrum of the SRF at each modulation frequency, the task of estimating $\{t_0, \lambda_0\}$ can only be accomplished if multiple frequency observations $z_{n\omega_0} = p_0^2 \hat{h}_{\text{FLI}}^*(n\omega_0)$, $n = 0, \dots, N-1$ are available. With $t_0 = 0$, estimation of λ_0 is a classical problem in system identification [28]. The presence of non-zero t_0 motivates development of new algorithms. In contrast to [21] where the authors use non-linear least squares fitting, here, we develop a closed form solution to the problem. In view of (14), let us set $y_n = z_{n\omega_0}/p_0^2$ and $\omega_n = n\omega_0$ and we have,

$$\frac{y_{n+1}}{y_n} = \frac{1 + j\omega_n \lambda_0}{1 + j\omega_{n+1} \lambda_0} e^{-j\omega_0 t_0} \quad \Leftrightarrow \quad y_{n+1} (1 + j\omega_{n+1} \lambda_0) = y_n (1 + j\omega_n \lambda_0) e^{-j\omega_0 t_0},$$

which is a linear system of equations in λ_0 and $e^{-j\omega_0 t_0}$ and can be solved with any four contiguous values such that $\ell \leq n < \ell + 4$, $\ell \in \mathbb{Z}$. We solve this linear system of equations for $\{y_n\}_{n=2}^{n=20}$ since the modulation frequencies in range $f_0 = 2 - 20$ MHz are highly stable. While system stability and signal integrity is one part that affects the SNR, variation of sample spread on the slide may also lead to weak emission. For this purpose, we use a confidence threshold for pixel amplitudes.

We compare our system identification approach with previously used non-linear least squares fitting [21] which uses $N = 40$ samples and favorable initialization conditions to obtain,

(nanoseconds)	$\tilde{\lambda}_0$	32.16	32.32	31.46	$\tilde{\lambda}_{0,\text{NNLS}}$	30.33	30.79	31.51
(meters)	$\tilde{d}$	2.54	2.49	2.54	$\tilde{d}_{\text{NNLS}}$	2.494	2.496	2.496

For comparable distance estimates, the identification method provides a more robust estimate of lifetimes. In comparison to NNLS fitting (based on Matlab's curve-fit toolbox), our relatively modest and non-iterative method is orders of magnitude faster ($\sim 10^3$). We close this subsection by highlighting that the phase/depth images in ToF sensors may be used for encoding interesting physical/material properties.

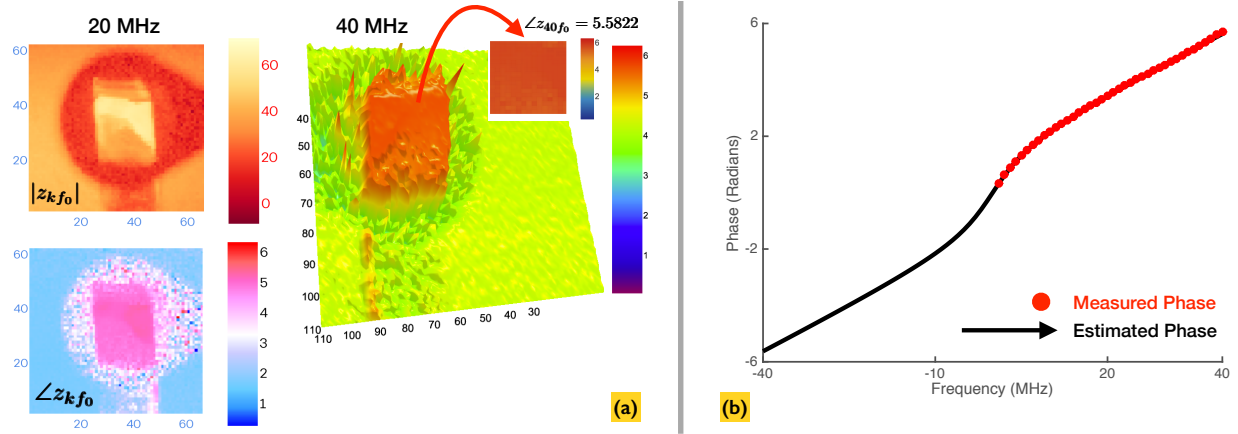


Fig. 7: Frequency Domain Time-of-Flight Imaging. (a) Multi-frequency measurements of scene containing a fluorescent sample with lifetime $\tau = 32$ ns. We show amplitude image $|z_{kf_0}|$ (in dB) as well as the phase image $\angle z_{kf_0}$ (in radians) with $f_0 = 1$ MHz. We also show a 3D plot of the phase image. This shows the effect of multipath interference. The phase at the background pixel is 4.1625 rads, which relates to a depth of 2.4842 m while the actual depth is 2.5 m. At this depth, the phase recorded as the sample location is 5.5822. The extra 3.3316 is attributed to the fluorescent lifetime decay phenomenon $\tan^{-1}(\omega\lambda_0)$. (b) At the sample location, we plot phase $\angle z_{kf_0} = -\angle \hat{h}(kf_0)$, $k = 1, \dots, 40$ (14) recorded by the ToF sensor. We also plot the phase estimated by our method.

Key takeaways of this section:

- ▶ Multi-path interference (MPI) in FD-ToF mode can be reinterpreted as a spectrum estimation problem. In order to recover MPI components, one must acquire multiple frequency measurements. Our demonstrations were based on matrix pencil method [26] but any other method is an option. For example, in [25], the authors use orthogonal matching pursuit.
- ▶ Transient imaging in TD-ToF mode can be recast as a sparse sampling problem [27] where the sampling kernel ϕ is the auto-correlation of the probing function. Whenever the probing function admits a bandlimited approximation, one may use spectral estimation to recover the transient image components.
- ▶ We demonstrated that fluorescence lifetime imaging can be performed with ToF sensors. We used system identification methods to estimate an unknown, parametric, transfer function linked with lifetime imaging.

IV. CONCLUSIONS AND FUTURE DIRECTIONS

We hope to have convinced the readers that ToF sensors are more than just depth sensors! In what follows, we make our concluding remarks with hints on possible future research directions.

ToF Imaging Pipeline We started with an image formation model that allows for studying different ToF modalities under one common framework. Almost all ToF systems can be characterized by the model parameters $\{p, h, \varphi\}$. Depending on the problem at hand, the role of $\{p, h, \varphi\}$ and the associated algorithms needs to be adapted. For example, the SRF discussed in context of MPI appears naturally in other problems such as ultrasound tomography [29] and light detection and ranging (LiDAR) [30]. However, the probing function and the IRF are very different for each case. In particular, consider the case of LiDAR. The probing function is modeled as $p = \delta$ and the IRF is a parametric function of form $\varphi(t) = \alpha e^{(a_k - T_0)t + b_k}$ where $\{a_k, b_k\}_{k=1}^4$ take 4 different values with continuous transitions, depending whether $t \in \{I_k\}_{k=1}^4$, where I_k is an instrument or sensor dependent quantity. This gives

rise to a new form of sampling kernel $\phi = \varphi$, as opposed to $\phi = p * \bar{p}$ (TD-ToF case). Hence, we believe that by systematically studying the role of $\{p, h, \varphi\}$ across various ToF problems—optical and non-optical—better insights may developed.

Another interesting direction may be to consider the case when the SRF is modeled by a differential equation. For example, in case of fluorescence lifetime imaging, the associated differential equation is $\mathcal{L}_\lambda = \partial_t + (1/\lambda)$ and the resulting reflected signal is the solution to $\Gamma \mathcal{L}_\lambda [r(t)] = p(t), t \geq 0$. The SRF in this case is the Green's function. Similar ideas may be used to develop algorithms for imaging through scattering/diffusive media where $\mathcal{L}$ is some differential equation that models diffusion. The parameters of $\mathcal{L}$ encode physical properties such as lifetime or scattering coefficient.

Probing Function Since the probing function is the only available degree-of-freedom in ToF imaging pipeline, it is important to understand what mathematical principles should be used for designing probing functions. Waveform design is a known art in radar and wireless communications. However, such options are rarely considered in optical ToF systems. Maximum length sequences for TD-ToF and sinusoids for the FD-ToF are the de facto examples. On the hand, it may not always be feasible to calibrate the probing function. In that case, it may be worthwhile to use blind deconvolution algorithms for image reconstruction.

Algorithms and Fundamental Limits MPI is a significant problem in ToF imaging and a number of papers have attempted to address this issue—both in time- and in frequency domain. However, till date, most of the results remain empirical and rarely discuss any details on performance guarantees. In a recent work [24], we used Cramér–Rao bounds in context of TD-ToF based multiple depth imaging. For instance, for the TD-ToF case, one may write the probing function as a Fourier Series with Fourier coefficients $\{\phi_m\}_{m \in \mathbb{Z}}$. In this case two optical paths, Δd apart, are resolvable provided that,

$$\frac{\Delta d}{T} \geq \frac{c}{4\pi} \frac{1}{\sqrt{\text{PSNR}}} \sqrt{\frac{S}{N}},$$

where $T = 2\pi/\omega_0$ is the fundamental time-period, PSNR is the peak-signal-to-noise-ratio, $S^{-1} = \sum_{m \in \mathbb{Z}} m^2 |\phi_m|^2$ and N is the number of measurements. For parametric SRFs, we believe that more such efforts could lead to hints on interesting applications of ToF sensors and motivate new problems in sampling theory. This could be the key to questions such as: when can two lifetimes in fluorescence lifetime imaging setup be super-resolved?

Modeling Non-ideal Reflections In our experience, the SRF of form $\Gamma_k \delta(t - t_k)$ (c. f. (8)) only approximately models a reflection. In a practical setting, $\Gamma_k \psi_k(t - t_k)$ may serve as a good starting point for modeling reflections. Here, ψ_k is a filter that models the interaction of the probing function with

the material property or accounts for distortion, system non-linearities and dispersive media. In seismic engineering, terahertz spectroscopy and ultrasound systems, this behavior is much more pronounced as material properties play an important role when the probing function undergoes a reflection.

FD-ToF This is specifically interesting mode of operation since most consumer ToF systems are based on FD-ToF which uses phase estimation. As seen in (6), if $\theta_\omega > 2\pi$, the depth estimates suffer with ambiguity or phase-wrapping problem. Previous solutions use co-prime frequencies, however, there is room for improvement. For example, in theory, phase is a linear function of frequency, $\theta_\omega = 2\omega d/c$. This is not the case in practice and leads to erroneous depth estimates with multiple frequency measurements. Hence, a desirable phase estimation algorithm should jointly correct for any distortions and phase-wrapping. Calling $m(t, \omega) \stackrel{(6)}{=} \Gamma_0 \left(1 + \frac{p_0^2}{2} \cos(\omega t + \theta_\omega)\right)$, θ_ω is estimated by sampling in time-domain, that is, $m_k = m(kT, \omega_0)$ given a fixed modulation frequency ω_0 . Alternatively, one may use multi-frequency sampling using $m_k = m(t_0, k\omega_0)$ for estimation of θ_ω . This gives rise to a broader question of when can time-frequency sampling be used, that, $m_{\ell,k} = m(\ell t_0, k\omega_0)$ in context of solving inverse-problems linked with ToF imaging. Specifically, when $\omega = t$ for depth imaging, the problem boils down to parameter estimation of chirp signals. Finally, in view of (10), *phase retrieval* algorithms may be designed when only intensity measurements are available, that is, $|m_k|^2 = |m(t_0, k\omega_0)|^2$.

Sensor Design for Higher Modulation Frequencies Most consumer-grade ToF sensors are based on continuous wave probing functions. Currently, such sensors work with high fidelity up till modulation frequency of about 80 MHz. We believe that a lot of the interesting physical phenomenon may only be observed as higher frequencies. For example, higher modulation frequencies will certainly enhance depth resolution and MPI correction capabilities. In context of fluorescence lifetime imaging [21], shorter lifetimes may be resolved with higher modulation frequencies. Similarly, sub-surface scattering properties can be studied with streak tubes [10]. This hints that indeed higher modulation frequencies are the pathway to scattered imaging. Such examples motivate the necessity of hardware or computational imaging solutions that can overcome the current technological limits.

ACKNOWLEDGEMENTS

AB thanks Prof. T. Blu for a first hand introduction to the topic of sparse sampling as well as mathematical advise over the course of years. In fact, a more general result linked with the problem of lifetime imaging was devised with the help of Prof. Blu's advise. The authors thank Chris Barsi, Refael Whyte, Jamie Schiel, Achuta Kadambi and Shahram Izadi for collaboration and various discussions pivoted around the experimental aspects of ToF imaging.

REFERENCES

- [1] O. Cossairt, M. Gupta, and S. K. Nayar, “When does computational imaging improve performance?” *IEEE Transactions on Image Processing*, vol. 22, no. 2, pp. 447–458, 2013.
- [2] R. Raskar, A. Agrawal, and J. Tumblin, “Coded exposure photography: motion deblurring using fluttered shutter,” *ACM Transactions on Graphics (TOG)*, vol. 25, no. 3, pp. 795–804, 2006.
- [3] S. S. Gorthi, D. Schaak, and E. Schonbrun, “Fluorescence imaging of flowing cells using a temporally coded excitation,” *Optics express*, vol. 21, no. 4, pp. 5164–5170, 2013.
- [4] P. E. Debevec and J. Malik, “Recovering high dynamic range radiance maps from photographs,” in *ACM SIGGRAPH 2008*. ACM, 2008, p. 31.
- [5] R. Ng, “Fourier slice photography,” in *ACM Transactions on Graphics*, vol. 24, no. 3. ACM, 2005, pp. 735–744.
- [6] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin, “Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing,” *ACM Transactions on Graphics*, vol. 26, no. 3, p. 69, 2007.
- [7] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. E. Kelly, and R. G. Baraniuk, “Single-pixel imaging via compressive sampling,” *IEEE Signal Processing Magazine*, vol. 25, no. 2, p. 83, 2008.
- [8] G. Zheng, R. Horstmeyer, and C. Yang, “Wide-field, high-resolution fourier ptychographic microscopy,” *Nature photonics*, vol. 7, no. 9, pp. 739–745, 2013.
- [9] R. I. Hartley and P. Sturm, “Triangulation,” *Computer vision and image understanding*, vol. 68, no. 2, pp. 146–157, 1997.
- [10] A. Velten, T. Willwacher, O. Gupta, A. Veeraraghavan, M. Bawendi, and R. Raskar, “Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging,” *Nature Comm.*, vol. 3, p. 745, 2012.
- [11] O. Katz, P. Heidmann, M. Fink, and S. Gigan, “Non-invasive single-shot imaging through scattering layers and around corners via speckle correlations,” *Nature Photonics*, vol. 8, no. 10, pp. 784–790, 2014.
- [12] M. Laurenzis, J. Klein, E. Bacher, and N. Metzger, “Multiple-return single-photon counting of light in flight and sensing of non-line-of-sight objects at shortwave infrared wavelengths,” *Optics letters*, vol. 40, no. 20, pp. 4815–4818, 2015.
- [13] G. Gariepy, N. Krstajić, R. Henderson, C. Li, R. R. Thomson, G. S. Buller, B. Heshmat, R. Raskar, J. Leach, and D. Faccio, “Single-photon sensitive light-in-flight imaging,” *Nature Communications*, vol. 6, 2015.
- [14] A. Kirmani, D. Venkatraman, D. Shin, A. Colaço, F. N. Wong, J. H. Shapiro, and V. K. Goyal, “First-photon imaging,” *Science*, vol. 343, no. 6166, pp. 58–61, 2014.
- [15] F. Heide, M. B. Hullin, J. Gregson, and W. Heidrich, “Low-budget transient imaging using photonic mixer devices,” *ACM Trans. Graph. (Proc. SIGGRAPH 2013)*, vol. 32, no. 4, pp. 45:1–45:10, 2013.
- [16] A. Kadambi, R. Whyte, A. Bhandari, L. Streeter, C. Barsi, A. Dorrington, and R. Raskar, “Coded time of flight cameras: sparse deconvolution to address multipath interference and recover time profiles,” *ACM Transactions on Graphics (TOG)*, vol. 32, no. 6, p. 167, 2013.
- [17] F. Heide, L. Xiao, W. Heidrich, and M. B. Hullin, “Diffuse mirrors: 3D reconstruction from diffuse indirect illumination using inexpensive time-of-flight sensors,” in *Proc. of IEEE CVPR*, Jun. 2014, pp. 3222–3229.
- [18] F. Heide, L. Xiao, A. Kolb, M. B. Hullin, and W. Heidrich, “Imaging in scattering media using correlation image sensors and sparse convolutional coding,” *Opt. Exp.*, vol. 22, no. 21, pp. 26 338–26 350, Oct 2014.
- [19] A. Kadambi, A. Bhandari, R. Whyte, A. Dorrington, and R. Raskar, “Demultiplexing illumination via low cost sensing and nanosecond coding,” in *Proc. of IEEE ICCP*, 2014, pp. 1–10.
- [20] N. Kitsunezaki, E. Adachi, T. Masuda, and J.-i. Mizusawa, “KINECT applications for the physical rehabilitation,” in *Proc. of IEEE MeMea*, 2013, pp. 294–299.

- [21] A. Bhandari, C. Barsi, and R. Raskar, "Blind and reference-free fluorescence lifetime estimation via consumer time-of-flight sensors," *Optica*, vol. 2, no. 11, pp. 965–973, Nov 2015.
- [22] S. Foix, G. Alenya, and C. Torras, "Lock-in time-of-flight (ToF) cameras: a survey," *IEEE Sensors Journal*, vol. 11, no. 9, pp. 1917–1926, 2011.
- [23] B. Büttgen, E. Mechat, M. Hamed-Ali, F. Lustenberger, and P. Seitz, "Pseudonoise optical modulation for real-time 3-d imaging with minimum interference," *IEEE Trans. on Circ. and Sys. I: Regular Papers*, vol. 54, no. 10, pp. 2109–2119, 2007.
- [24] A. Bhandari, A. Kadambi, and R. Raskar, "Sparse linear operator identification without sparse regularization? Applications to mixed pixel problem in tof/range imaging," in *Proc. of IEEE ICASSP*, 2014, pp. 365–369.
- [25] A. Bhandari, A. Kadambi, R. Whyte, C. Barsi, M. Feigin, A. Dorrington, and R. Raskar, "Resolving multipath interference in time-of-flight imaging via modulation frequency diversity and sparse regularization," *Optics Letters*, vol. 39, no. 7, 2014.
- [26] Y. Hua, A. Gershman, and Q. Cheng, *High-resolution and robust signal processing*. CRC Press, 2003.
- [27] T. Blu, P.-L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot, "Sparse sampling of signal innovations," *IEEE Signal Proc. Mag.*, vol. 25, no. 2, pp. 31–40, 2008.
- [28] L. Ljung, *System identification*. Springer, 1998.
- [29] R. Tur, Y. C. Eldar, and Z. Friedman, "Innovation rate sampling of pulse streams with application to ultrasound imaging," *IEEE Trans. Sig. Proc.*, vol. 59, no. 4, pp. 1827–1842, 2011.
- [30] S. Hernandez-Marin, A. M. Wallace, and G. J. Gibson, "Bayesian analysis of lidar signals with multiple returns," *IEEE Trans. on Patt. Anal. and Machine Intell.*, vol. 29, no. 12, pp. 2170–2180, 2007.

AUTHORS

A. Corresponding Author (Contact Details and Bio)

Ayush Bhandari

E14-374 H, Massachusetts Institute of Technology,
75 Amherst Street, Cambridge, MA 02139-4307 USA.

Email: ayush@MIT.edu

Phone: +1 617 682 9874

Ayush Bhandari is currently a researcher at the MIT. Before moving to MIT, he has held various research positions starting with INRIA–Rennes, France followed by NTU Singapore and then, CUHK Hong Kong with the most recent appointment at EPFL Switzerland. His research interests include sampling and approximation theory and, inverse problems linked with computational imaging. He was recipient of the Best Student Paper Award at IEEE International Symposium on Signal Processing and Its Applications (2007) and the Best Paper Award at the IEEE Systems and Information Engineering Design Symposium (2007).

B. Co-Author (Bio)

Ramesh Raskar heads the Camera Culture research group at the MIT Media Lab. His research interests span the fields of computational photography, inverse problems in imaging and human-computer interaction. Raskar has been the recipient of numerous awards including the TR100 Award from Technology Review, the Global Indus Technovator Award, Sloan Research Fellowship (2009) and the Darpa Young Faculty award (2010). Other awards include Marr Prize honorable mention 2009, LAUNCH Health Innovation Award and the Vodafone Wireless Innovation Project Award (first place), 2011. He holds over 40 US patents and is currently co-authoring a book on Computational Photography.