SIAM J. IMAGING SCIENCES
Vol. 18, No. 2, pp. 1439–1467

© 2025 Society for Industrial and Applied Mathematics

Blind Time-of-Flight Imaging: Sparse Deconvolution on the Continuum with Unknown Kernels*

Ruiming Guo† and Ayush Bhandari†

Abstract. In recent years, computational time-of-flight (ToF) imaging has emerged as an exciting and novel imaging modality that offers new and powerful interpretations of natural scenes, with applications extending to three-dimensional, light-in-flight, and non-line-of-sight imaging. Mathematically, ToF imaging relies on algorithmic super-resolution, as the back-scattered sparse light echoes lie on a finer time resolution than what digital devices can capture. Traditional methods necessitate knowledge of the emitted light pulses or kernels and employ sparse deconvolution to recover scenes. Unlike previous approaches, this paper introduces a novel, blind ToF imaging technique that does not require kernel calibration and recovers sparse spikes on a continuum, rather than a discrete grid. By studying the shared characteristics of various ToF modalities, we capitalize on the fact that most physical pulses approximately satisfy the Strang–Fix conditions from approximation theory. This leads to a new mathematical formulation for sparse super-resolution. Our recovery approach uses an optimization method that is pivoted on an alternating minimization strategy. We benchmark our blind ToF method against traditional kernel calibration methods, which serve as the baseline. Extensive hardware experiments across different ToF modalities demonstrate the algorithmic advantages, flexibility, and empirical robustness of our approach. We show that our work facilitates super-resolution in scenarios where distinguishing between closely spaced objects is challenging, while maintaining performance comparable to known kernel situations. Examples of light-in-flight imaging and light-sweep videos highlight the practical benefits of our blind super-resolution method in enhancing the understanding of natural scenes.

Key words. computational imaging, inverse problems, blind deconvolution, super-resolution, time-of-flight imaging

MSC codes. 62H35, 68U99, 78A46, 94A12

DOI. 10.1137/24M1657638

1. Introduction to time-of-flight imaging.

The emerging theme of *computational sensing and imaging* or CoSI [10, 16] has catalyzed never-seen-before capabilities in the context of imaging and vision. Some noteworthy examples are *single-pixel* imaging [25, 48], *non-line-of-sight* imaging [55, 60], *ultrafast* imaging at a trillion [55] and a billion [28] frames per second, and imaging of *black holes* [51]. In all such examples and beyond, the fundamental difference from the traditional viewpoint is the new mindset that altering the forward model of an

---

[*] Received by the editors April 29, 2024; accepted for publication (in revised form) October 28, 2024; published electronically May 30, 2025.

https://doi.org/10.1137/24M1657638

Funding: The work of the authors is supported by the UK Research and Innovation council's FLF Program "Sensing Beyond Barriers via Non-Linearities" (MRC Fellowship award MR/Y003926/1).

[†] Department of Electrical and Electronic Engineering, Imperial College London, London, SW72AZ, UK (ruiming.guo@imperial.ac.uk, a.bhandari@imperial.ac.uk, ayush@alum.mit.edu).

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

imaging system enables new capabilities—one can *co-design* hardware and algorithms in pursuit of new advantages that cannot be harnessed by optimizing hardware or algorithms alone.

In the context of CoSI, the hardware-software co-design approach has been particularly fruitful for time-of-flight (ToF) imaging. A notable characteristic of ToF imaging is that each pixel of the imaging sensor captures a scene-dependent time profile at a specific time resolution, usually on a scale ranging from nanoseconds (ns) to picoseconds (ps). This technique, therefore, is also referred to as time-resolved imaging, highlighting its ability to capture detailed temporal information within a scene (see Chapter 5 in [10]), as illustrated in Figure 1. In contrast to conventional imaging paradigms, ToF imaging offers numerous advantages that were previously unimaginable. This is due to the fact that each ToF measurement comprises two types of images: the conventional 2D photograph or amplitude image, and the unconventional time-resolved, depth image. The advent of ToF imaging technology paves the way for novel methods and applications in various domains, including computer vision [21, 34, 52], graphics [32, 35, 61], autonomous vehicles, and biomedical imaging [5, 46]. Two concrete examples that also serve as an experimental validation of our work (in section 5) are as follows.

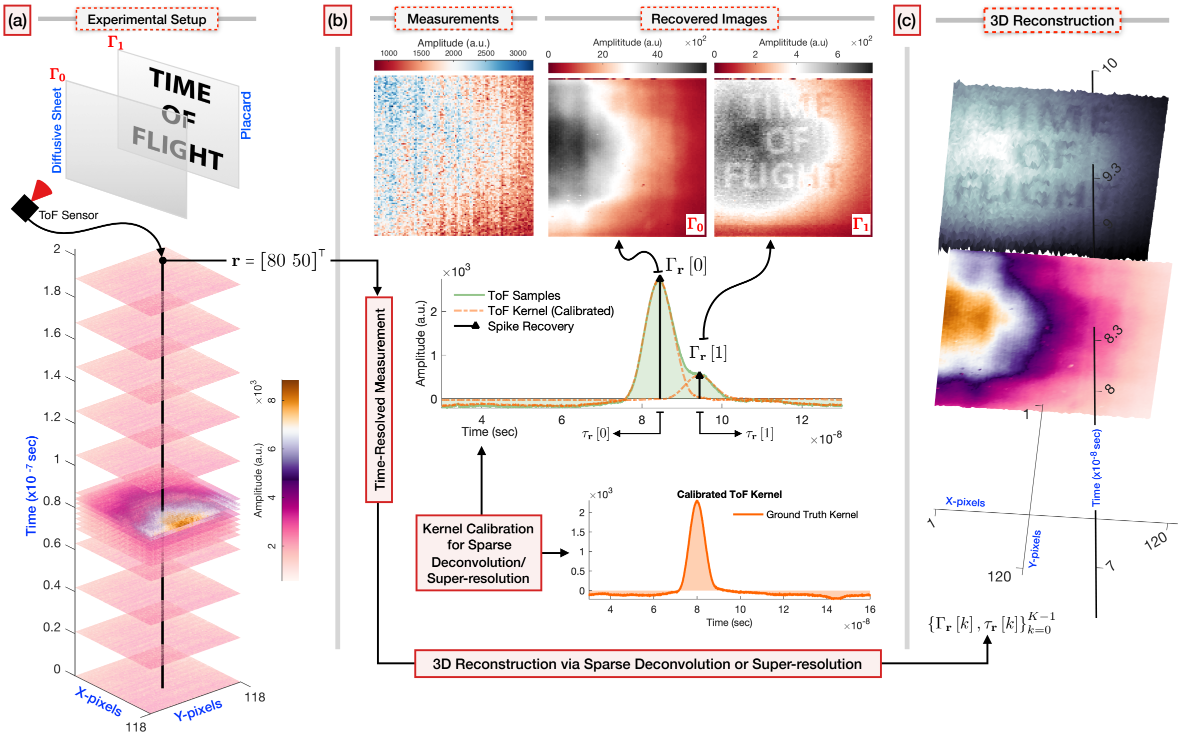


Figure 1. *Diffusive imaging.* (a) *Experimental setup* [13, 35]*. A "TIME OF FLIGHT" placard is hidden by a diffusive semitranslucent sheet between the placard and the ToF sensor.* (b) *Depth estimation from pixel measurements via sparse deconvolution techniques* [13, 14, 17]*. The ToF kernel (i.e.,* $\varphi_{\mathbf{r}}$ *in* (2.2)*) is calibrated and known prior to experiments.* (c) *Three-dimensional (3D) scene reconstruction from estimated depth parameters in* (b)*. The main goal of this paper is to recover a 3D scene without preknown kernel information (refer to Figures* 8 *and* 10*), aiming to eliminate the necessity for experimental calibration.*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

(1) Diffuse imaging [8, 13, 32, 35]. As illustrated in Figure 1, here ToF imaging enables recovery of objects hidden or obfuscated via diffusive materials; also see section 5.1.1.
(2) Light-in-flight (LIF) imaging [1, 29, 32, 35, 56]. In 1978 [1], Abramson pioneered LIF using a holographic method to record the wavefront of a light pulse as it traveled and scattered off a white screen. In a similar spirit, ToF measurements also enable LIF imaging that involves capturing and reconstructing the trajectory of light as it traverses and interacts with various objects in the scene, leading to completely new pathways for scene understanding. This is another ToF imaging capability, covered in this paper in section 5.2 (https://youtu.be/ffkc_z8ogE8).

Depending on the technology and temporal resolution, ToF imaging setups can be broadly classified into three categories: (i) lock-in sensors [27] operating at nanosecond resolution, (ii) single photon avalanche diode detectors [29, 37] operating at picosecond resolution, and (iii) streak tubes [55] operating at femtosecond resolution. Among them, lock-in sensors are notably popular for ToF imaging, largely because they are available as consumer-grade technology, exemplified by *photonic mixer devices* and the Microsoft Xbox One's Kinect, making them widely used and accessible.

1.1. Fundamental role of sparsity and super-resolution in ToF imaging.

In conventional 2D photography the scene is rich in its natural representation and described by low-level image features, such as edges, corners, and ridges [26]. The resulting images are sparse in transform domain, e.g., discrete cosine transform or wavelet basis. However, when the same scene is interpreted from the perspective time-scales [32, 35, 61], the features are fundamentally different and take the form of global and direct delays, interreflections, and subsurface scattering. Mathematically, this translates to the fact that scenes are naturally sparse along the temporal dimension. This is also clearly seen in Figure 1, where the time profile along a particular pixel comprises two Dirac impulses. Clearly, in terms of the Shannon–Nyquist method of digitization, acquiring such scenes would entail exorbitant sampling rates which may be technologically unviable, not to mention how expensive and challenging it may be to implement such systems. That said, realizing that ToF measurements encode sparse features, super-resolution (SRes) or the problem of recovering spikes from filtered kernels [17, 19, 23, 24, 42] turns out to be the most appropriate setup for ToF imaging methods [4]. The main advantage of resorting to such a flavor is that nonbandlimited objects such as Dirac impulses can be recovered from filtered measurements, without requiring a sampling criterion [13]. Hence, resorting to SRes formulation proves to be highly beneficial in lowering the hardware constraints.

1.2. Related work.

In the existing literature, the ToF imaging methods via SRes formulation can be broadly classified as follows. **Stochastic approaches.** The authors in [33] proposed a Bayesian framework to find out the 3D scene parameters via evaluating posterior probability distribution. Adam et al. [2] utilize Bayesian inference for the recovery of 3D scene geometry, such as shape, illumination, and albedo. **Optimization approaches.** Kadambi et al. in [35] formulated the time profile recovery with multipath interference as sparse deconvolution and utilized an OMP (orthogonal matching pursuit) algorithm to solve for spike estimation. In [32], Heide et al. posed the scene recovery as a temporal-spatial regularization problem, encompassing both spatial and temporal regularized penalties. An alternating minimization strategy was utilized to split the principle problem into two tractable subproblems

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

via proximal operators. **Parametric approaches.** Built on the sparse representation of ToF measurements, the authors in [9, 13, 14] first modeled the inverse problem of scene recovery as deconvolution, which essentially boils down to the spectral estimation problem [22]. Subsequently, Prony's method is employed to super-resolve the parameters of the 3D scene.

1.3. Motivation for blind ToF imaging. By comparing and contrasting the mathematical SRes [17, 19, 23, 24, 42] and ToF SRes [2, 13, 14, 32, 33, 35] communities, the key takeaways from the existing art can be distilled as follows.

1. Significance of kernel calibration: A vital assumption in any SRes technique [17, 19, 23, 24, 42] is the knowledge of the kernel. In the ToF context [2, 13, 14, 32, 33, 35], this translates to kernel calibration prior to conducting experiments (see Figure 1(b)). Since ToF systems employ active sensing—where the illumination source is separate from the imaging sensor—any alterations in illumination necessitate a recalibration of the kernel related to illumination. Over the illumination system's lifespan, physical factors like temperature and optical setup changes may cause variability in the kernel. From the perspective of solving an inverse problem, any discrepancy between the calibrated kernel and ToF measurements can significantly degrade the quality of spike estimation, thus limiting the adaptability and reliability of ToF imaging methods. This highlights the importance of developing blind ToF methods that operate without assuming the knowledge of the kernel [3].
2. Community divide: With the exception of parametric approaches [9, 13, 14], most ToF SRes methods do not yet leverage "off-the-grid" model for spikes or Dirac impulses. This means the unknown sparse signal is presumed to "live" on a discrete grid, an assumption that may not reflect reality. Conversely, a notable advancement within the mathematical SRes [17, 19, 23, 42] community has been the ability to recover spikes on the continuum. We believe this gap in understanding is largely due to the minimal interaction between the two fields. Moreover, state-of-the-art recovery techniques might not directly transfer to ToF imaging, as they are not tailored for handling large-scale measurements, which can extend to tensors. This calls for the creation of efficient algorithms capable of modeling Dirac impulses continuously and managing the voluminous data produced by ToF sensors. Crucially, the effectiveness of these approaches hinges on their validation and benchmarking through real-world experiments.

1.4. Contributions. The main goal of this paper is to develop a flexible, robust, and efficient blind ToF imaging method that allows for super-resolved 3D scene recovery in realistic experimental setups, independent of any kernel calibration. We outline the contributions of our work as follows:

- **Mathematical model:** Leveraging our experience with ToF imaging [10, 13], we propose a generic model for a ToF imaging pipeline in which the Dirac impulses[1] are modeled on the continuum.

  Different from previous works in computational imaging and computer vision/graphics, we assume that the ToF kernel is unknown and is equipped with Strang–Fix properties. This is to say, the shifts of the ToF kernel reproduce exponential-polynomial functions

---

[1] We consider continuous-time spikes in terms of distribution.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

(see (3.1)) [50]. This conceptualization is a key enabler for blind ToF imaging in that Strang–Fix conditions lead to

  (i) a flexible choice of unknown kernel, covering the well-known modalities, e.g., lock-in sensors [27] and time-correlated single photon counting (TCSPC) systems [40],
  (ii) a key continuous-time recovery of sparse spikes.

- **Recovery algorithm:** We design a robust, efficient, and scalable algorithm that achieves fill 3D-stack processing of ToF measurements. We resort to a nonconvex optimization scheme to estimate the unknown kernel and 3D scene parameters via an alternating minimization strategy, which we validate in various experimental setups and datasets.
- **Experimental validation and benchmarking:** Through a series of 10 experiments, we benchmark our proposed approach against existing methods with kernel calibration, which serves as a ground truth, thus showcasing the comparable performance in various real-world settings, such as
  (i) looking through diffusers (see 5.1.1),
  (ii) large-scale data processing ($120 \times 120 \times 3968 \times 4$ image tensor; see section 5.1.2),
  (iii) intertarget separation of 0.20 cm (see section 5.1.2),
  (iv) high-order imaging (see section 5.1.3), and
  (v) LIF imaging at nanosecond scale (see section 5.2).

These experiments corroborate the adaptability, robustness, and SRes capability of our method.

In the absence of kernel calibration, our setting takes the form of a blind sparse deconvolution (BSD) problem. In the prior art, BSD was studied in various flavors including BSD via multichannel [36, 58] and single-channel [38, 59] methods. However, such BSD approaches do not translate to the ToF context due to (1) kernel priors, e.g., incoherence [39], nonnegativity [41], and short-support [59], (2) the assumption that Dirac impulses or spikes lie on a grid [36, 45, 59], and (3) the large-scale data that arises in ToF imaging [13, 35]. Finally, we find that prevailing works on BSD are not designed to handle the ToF pipeline. The lack of experimental validation of such approaches not only creates a gap between theory and practice but also makes benchmarking of these approaches very challenging.

Notation. The sets of integer, real, and complex-valued numbers are denoted by $\mathbb{Z}, \mathbb{R}$, and $\mathbb{C}$, respectively. The set of N contiguous integers is denoted by $\mathbb{I}_N = \{0, \ldots, N-1\}, N \in \mathbb{Z}^+$. Continuous functions are written as $g(t), t \in \mathbb{R}$; their discrete counterparts are represented by $g[n] = .g(t)|_{t=nT}$, $n \in \mathbb{Z}$, where $T > 0$ takes the role of sampling period. Vectors and matrices are written in bold lowercase and uppercase fonts, such as $\mathbf{g} = [g[0], \ldots, g[N-1]]^\top \in \mathbb{R}^N$ and $\mathbf{G} = [g_{n,m}]_{n \in \mathbb{I}_N}^{m \in \mathbb{I}_M} \in \mathbb{R}^{N \times M}$. The $\mathbf{L}_p(\mathbb{R})$ space equipped with the p-norm or $\|\cdot\|_{\mathbf{L}_p(\mathbb{R})}$ is the standard Lebesgue space. For instance, $\mathbf{L}_1$ and $\mathbf{L}_2$ denote the space of absolute and square-integrable functions, respectively. Spaces associated with sequences are denoted by ℓ_p. The max-norm ($\mathbf{L}_\infty$) of a function is defined as $\|g\|_\infty = \inf\{c_0 \geqslant 0 : |g(t)| \leqslant c_0\}$; for sequences, we use, $\|g\|_\infty = \max_n |g[n]|$. The $\mathbf{L}_2$-norm of a function is defined as $\|g\|_2 = \sqrt{\int |g(t)|^2 dt}$, while for sequences, we have $\|g\|_2 = \sqrt{\sum_{n=0}^{N-1} |g[n]|^2}$. The ℓ_0-norm of a sequence denotes the cardinality (i.e., the number of nonzero entries) of the sequence. The inner-product of two functions $f, g \in \mathbf{L}_2$ is defined as $\langle f, g \rangle = \int f(t) g^*(t) dt$, while for sequences, we have $\langle f, g \rangle = \sum_{n=0}^{N-1} f[n] g^*[n]$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

The vector space of polynomials with complex coefficients and degrees less than or equal to K is denoted by P_K, for instance, $\mathsf{Q}_{\mathbf{r}}(z) = \sum_{k=0}^{K} h_k z^k \in P_K$. The N-order derivative of a function is denoted by $\partial_t^{(N)} g(t)$. The space of first-order continuously differentiable, real-valued functions is denoted by $C(\mathbb{R})$. For sequences, the first-order finite difference is denoted by $(\Delta g)[n] = g[n+1] - g[n]$. For any exponential type functions $|g(t)| \leqslant Ae^{B|t|}$, its Laplace transform is defined by $\mathcal{L}_g(s) = \int_0^\infty g(t)e^{-st}dt, s \in \mathbb{C}$. For any function $g \in \mathbf{L}_1$, its Fourier transform is defined by $\widehat{g}(\omega) = \int g(t)e^{-\jmath\omega t}dt$. For sequences, the discrete Fourier transform (DFT) of a sequence $\mathbf{g} \in \ell_1$ is denoted by $\widehat{g}[m] = \sum_{n=0}^{N-1} g[n]e^{-\jmath\frac{2\pi n}{N}m}$. Let $\mathbf{W}_N^M$ and $\mathbf{V}_N^M$ denote the $N \times M$ Vandermonde matrices $\mathbf{W}_N^M = [\xi_N^{-n\cdot m}]_{n\in\mathbb{I}_N}^{m\in\mathbb{I}_M}$, $\mathbf{V}_N^M = \frac{1}{N}[\xi_N^{n\cdot m}]_{n\in\mathbb{I}_N}^{m\in\mathbb{I}_M}$, $\xi_N^n = e^{\jmath\frac{2\pi n}{N}}$. The DFT of $\mathbf{g}$ can be expressed as $\widehat{\mathbf{g}} = \mathbf{W}_N^N \mathbf{g}$. The Gram matrix of $\mathbf{G}$ is defined as $\mathcal{G}(\mathbf{G}) = \mathbf{G}^{\mathsf{H}}\mathbf{G}$. The shorthand notation for diagonal matrices is given by $\mathscr{D}_K(\mathbf{h})$ with $[\mathscr{D}_K(\mathbf{h})]_{k,k} = [\mathbf{h}]_{k\in\mathbb{I}_K}$. The mean-squared error (MSE) between $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N$ is given by $\mathcal{E}(\mathbf{x},\mathbf{y}) = \frac{1}{N}\sum_{n=0}^{N-1} |x[n] - y[n]|^2$.

2. Image formation. Let $\mathbf{r} = [x, y]^\top$ denote a point in the Cartesian coordinate where $(\cdot)^\top$ is the transpose operation. ToF sensors are active imaging systems that probe the 3D scene of interest with some time-localized kernel [46, 49, 55, 57], denoted by $p_{\mathbf{r}}(t)$ at a point $\mathbf{r}$. Based on the choice of kernel, ToF imaging can be classified as time-domain and frequency-domain ToF setups.[2]

The emitted signal $p_{\mathbf{r}}(t)$ interacts with the 3D scene characterized by the spatio-temporal scene response function (SRF) $s_{\mathbf{r}}(t,t')$. In the case of multiple reflections (see Figure 2), the SRF is given by

$$s_{\mathbf{r}}\left(t,t'\right) = \sum_{k=0}^{K-1} \Gamma_{\mathbf{r}}\left[k\right] \delta\left(t - t' - \tau_{\mathbf{r}}\left[k\right]\right), \tag{2.1}$$

where $\delta(\cdot)$ is a Dirac distribution and $\{\Gamma_{\mathbf{r}}[k], \tau_{\mathbf{r}}[k]\}_{k\in\mathbb{I}_K}$ are the corresponding reflectivities and time-delays ($\tau_{\mathbf{r}}[k] = 2d_{\mathbf{r}}[k]/c$) induced by K light paths at point $\mathbf{r}$.[3]

In several practical scenarios, the SRF can be written as a shift-invariant function, $s_{\mathbf{r}}(t,t') = s_{\mathbf{r}}(t-t')$, which reduces the measurements to a convolution format. This is because the interplay of the emitted signal with the 3D scene results in the reflected signal given by $r_{\mathbf{r}}(t) = \int p_{\mathbf{r}}(t')s_{\mathbf{r}}(t,t')dt'$ resulting in convolution, $r_{\mathbf{r}}(t) = (p_{\mathbf{r}} * s_{\mathbf{r}})(t)$. The reflected signal is measured at the ToF sensor through its electro-optical architecture, which is described by its instrument response function (IRF), denoted by $\psi_{\mathbf{r}}(t,t')$.[4] This leads to the continuous-time

---

[2] While this paper is pivoted around time-localized kernels, to keep the exposition as general as possible, we will adhere to the generalized model described in our series of papers [5, 6, 7, 13, 14]. This model is general in the sense that it not only consolidates both time and frequency domain approaches but is also compatible with the broader theme of the ToF principle used in areas like terahertz [44], ultrasound [53], and seismic imaging [20] as well as optical coherence tomography [15, 47] and LIDAR [18].

[3] Note the SRF may be described more generally where $s_{\mathbf{r}}(t,t')$ is a Green's function of some partial differential equation [5] or simply a transfer function. For instance, in the scenario of fluorescence lifetime imaging, $s_{\mathbf{r}}(t,t') = s_{\mathbf{r}}^{\mathrm{Depth}}(t,t') + s_{\mathbf{r}}^{\mathrm{Decay}}(t,t')$ for which $s_{\mathbf{r}}^{\mathrm{Depth}}(t,t')$ is defined in (2.1) and represents a delay of $\tau_{\mathbf{r}}[k]$ due to the fluorescent sample's placement at depth $d_{\mathbf{r}}[k]$ meters from the sensor, and $s_{\mathbf{r}}^{\mathrm{Decay}}(t,t') = \alpha_{\mathbf{r}}\exp(-(t-t'-\tau_{\mathbf{r}})/\lambda)h(t-t'-\tau_{\mathbf{r}})$ where $\alpha_{\mathbf{r}}$ and λ are emission light strength and lifetime of the fluorescent sample, respectively, and $h(t)$ is the Heaviside step function.

[4] In the absence of time-resolved perspective or steady-state scene assumption, the IRF is equivalent to the spatial point spread function.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

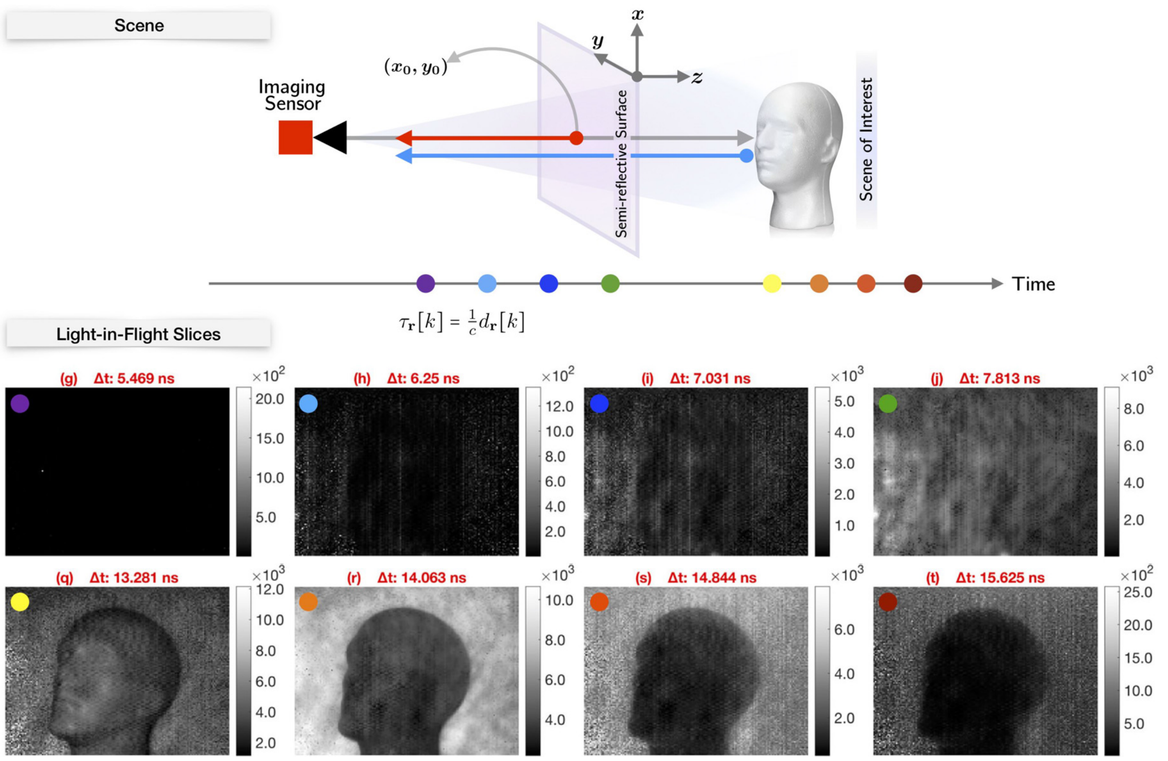


Figure 2. *Pipeline for LIF imaging (see https://youtu.be/ffkc_z8ogE8). Top row: Experimental setup capturing light's interaction with a scene, consisting of a mannequin head positioned between a diffusive surface and a backdrop wall. Bottom row: Using a ToF camera, we visualize the light's progression across different time instances marked by "•," highlighted by visible light's spectrum colors coding the ascending time sequence. Initially, light encounters the diffusive sheet in the early time-slots,* i.e., (i) *(blue) and* (j) *(green). Light then moves over the mannequin and ultimately reaches the back wall, as shown in frames* (q) *(yellow) to* (t) *(red). The LIF accurately maps the* 3*D geometry of the scene. The full LIF time slices are shown in Figure* 14.

measurements defined by $g_{\mathbf{r}}(t) = \int r_{\mathbf{r}}(t')\psi_{\mathbf{r}}(t,t')dt'$. Again, in practice, the IRF $\psi_{\mathbf{r}}(t,t')$ is typically shift-invariant, i.e., $\psi_{\mathbf{r}}(t,t') = \psi_{\mathbf{r}}(t-t')$. Consequently, the measurements impinging on the ToF imaging sensor simplify to $g_{\mathbf{r}}(t) = (r_{\mathbf{r}} * \psi_{\mathbf{r}})(t)$, which can be further simplified as

$$g_{\mathbf{r}}(t) = (s_{\mathbf{r}} * \varphi_{\mathbf{r}})(t) \quad \text{and} \quad \varphi_{\mathbf{r}}(t) \overset{\text{def}}{=} (p_{\mathbf{r}} * \psi_{\mathbf{r}})(t), \tag{2.2}$$

where $\varphi_{\mathbf{r}}$ can be interpreted as the kernel in the spirit of SRes [17] or the signal processing perspective. Finally, uniform sampling of the continuous-time input results in per-pixel digital measurements,

$$g_{\mathbf{r}}[n] = g_{\mathbf{r}}(t)|_{t=nT} = \sum_{k=0}^{K-1} \Gamma_{\mathbf{r}}[k]\, \varphi_{\mathbf{r}}(nT - \tau_{\mathbf{r}}[k]), \tag{2.3}$$

where $T > 0$ is the sampling step and usually of the magnitude of 60 to 100 picoseconds (see section 5). A global view of the image formation process is described as a mathematical block diagram in Figure 3.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

$$\underbrace{p_{\mathbf{r}}}_{\text{Emitted Signal}} \rightarrow \underbrace{\boxed{\boxed{s_{\mathbf{r}}}}}_{\text{SRF}} \rightarrow \underbrace{r_{\mathbf{r}}(t) = \int p_{\mathbf{r}}\left(t'\right) s_{\mathbf{r}}\left(t,t'\right) dt'}_{\text{Reflected Signal}} \rightarrow \underbrace{\boxed{\boxed{\psi}}}_{\text{IRF}} \rightarrow \underbrace{g_{\mathbf{r}}(t) = \int r_{\mathbf{r}}\left(t'\right) \psi_{\mathbf{r}}\left(t,t'\right) dt'}_{\text{Measured Signal}} \xrightarrow{\text{Sampling}} \underbrace{g_{\mathbf{r}}[n] = g_{\mathbf{r}}(t)|_{t=nT}}_{\text{Measured Samples}}$$

Figure 3. *Block diagram for ToF image formation process. The goal is to estimate* $s_{\mathbf{r}}$ *from* $\{g_{\mathbf{r}}[n]\}_{n\in\mathbb{I}_N}$.

Goal. Starting with $\{g_{\mathbf{r}}[n]\}_{n\in\mathbb{I}_N}$ *only*, our goal is to recover $s_{\mathbf{r}}(t)$ and $\varphi_{\mathbf{r}}(t)$ from possibly imperfect and distorted measurements.

3. Super-resolved ToF imaging with Strang–Fix properties.

Due to the shift-invariant property of the SRF and IRF, the ToF measurements depend on the kernel $\varphi_{\mathbf{r}}(t) \overset{(2.2)}{=} (p_{\mathbf{r}} * \psi_{\mathbf{r}})(t)$. The literature around the topic of ToF imaging is largely focused on the data model with discrete sparsity so that sparse recovery techniques can be leveraged [32, 35, 43, 52]. However, such approaches artificially constrain delays on a discrete grid, i.e., $\{\tau_{\mathbf{r}}[k]\}_{k\in\mathbb{I}_K} \subseteq T\mathbb{Z}$, which is an artifact of the data model. We believe that this is the reason why SRes methods have not been widely reported in the ToF literature.

In developing a blind ToF imaging strategy, an effective starting point involves abstracting properties of $\varphi_{\mathbf{r}}(t)$ (2.2) common across various ToF modalities [5, 6, 7, 13, 14, 32, 35, 40, 52], such as lock-in sensors and TCSPC systems. Our observation, based on various experimentally calibrated kernels used in ToF imaging [5, 6, 7, 13, 14, 32, 35, 40, 52], indicates that these kernels exhibit shared characteristics. By identifying and abstracting these common mathematical properties, we can pave the way for blind ToF imaging strategies. Specifically, our analysis of the literature highlights these shared features, setting the foundation for our approach.

- *Time-localization*: $\varphi_{\mathbf{r}}(t) = 0,\ t \notin \mathcal{D} \subset \mathbb{R}$, where $\mathcal{D}$ is a finite, contiguous interval in the time domain.
- *Smoothness*: $\varphi_{\mathbf{r}} \in C(\mathbb{R})$ is continuously differentiable.

Toward kernels with Strang–Fix properties. Given the nature of ToF kernels, a weighted linear combination of the kernel reproduces the prototype function,

$$t^h e^{-\frac{\nu_m t}{T}}, \quad h \in \mathbb{N} \text{ and } \nu_m \in \mathbb{C},$$

and hence, implicit parameterization in terms of $t^h e^{-\frac{\nu_m t}{T}}$ captures the features of the kernel in a *flexible yet compressible manner.* As we shall see shortly, this compressibility proves to be advantageous in our blind ToF context. Similar properties have been widely studied in the wavelet and approximation theory, where they arise in the context of *Strang–Fix conditions* [50, 54]. In particular, such kernels satisfy the property

$$t^h e^{-\frac{\nu_m t}{T}} = \sum_{n\in\mathbb{Z}} \mu_{m,n,h} \overline{\varphi}_{\mathbf{r}}(t - nT),\ m \in \mathbb{I}_M \text{ and } h \in \mathbb{I}_H, \tag{3.1}$$

where $\overline{\varphi}_{\mathbf{r}}(t) = \varphi_{\mathbf{r}}(-t)$ and $\{\mu_{m,n,h}\}_{n\in\mathbb{Z}}$ are the exponential-polynomial reproducing coefficients. A visual illustration of (3.1) is shown in Figure 4. In other words, $\{t^h e^{-\nu_m \frac{t}{T}}\}_{m\in\mathbb{I}_M}$ belongs to a shift-invariant space spanned by shifts of $\varphi_{\mathbf{r}}$. For this to work, the kernel $\varphi_{\mathbf{r}}$ must satisfy what is known as the generalized Strang–Fix conditions [54]:

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

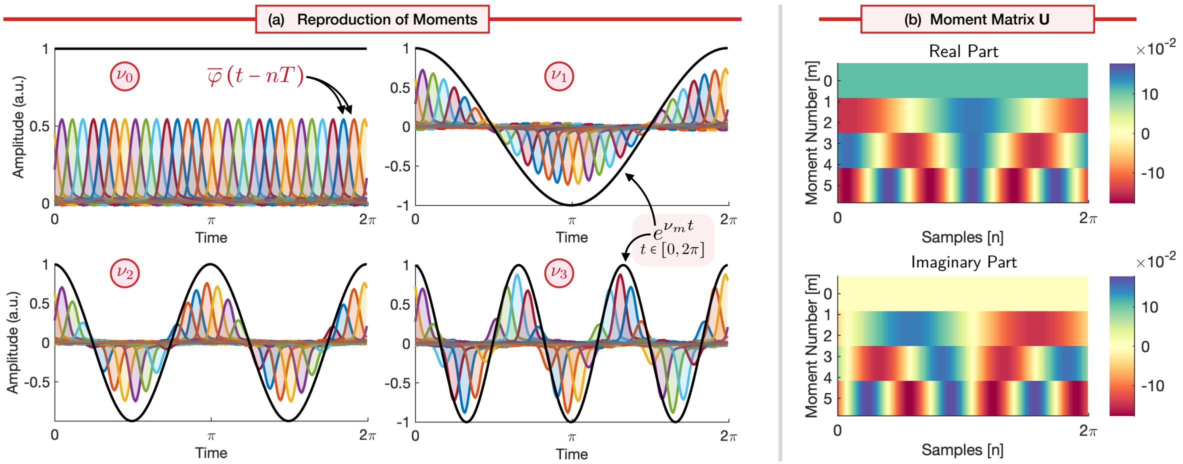


Figure 4. *Using the kernel* $\varphi_{\mathbf{r}}$ *in Figure* 8 (a), *we reproduce real part of exponentials* $e^{\nu_m t}$ *in* (a) $m=0,\ldots,3$. (b) *Visualization of the moment coefficient matrix* $\mathbf{U}$ *in* (3.9).

$$\partial_\nu^{(h)}\mathcal{L}_{\varphi_{\mathbf{r}}}(\nu)\Big|_{\nu=\nu_m}\neq 0 \quad\text{and}\quad \partial_\nu^{(h)}\mathcal{L}_{\varphi_{\mathbf{r}}}(\nu)\Big|_{\nu=\nu_m+\jmath\frac{2\pi l}{T}}=0,\ l\in\mathbb{Z}/\{0\},\ m\in\mathbb{I}_M \text{ and } h\in\mathbb{I}_H, \tag{3.2}$$

where $\mathcal{L}_{\varphi_{\mathbf{r}}}(s)$ denotes the Laplace transform of $\varphi_{\mathbf{r}}$ and M in (3.1) is the order of the kernel.

The main advantage of the parametrization in (3.1) is that the continuous-time unknown kernel ($\varphi_{\mathbf{r}}(t)$) can be expressed in a low-dimensional, discrete representation, $\{\mu_{m,n,h}\}_{n\in\mathbb{Z}}$ which makes the formulation of our optimization method efficient and tractable.

Perfect recovery of SRF with known kernel. Before developing blind ToF imaging methodology, we consider the intermediate step where the kernel is assumed to be known. In this setting, we consider the following two fundamental questions: (1) What is the mathematical criterion for perfect recovery of the SRF? and (2) Is there a constructive algorithm for recovery? In what follows, we will demonstrate an SRF recovery method that relies on transform domain characterization of spikes and leverages spectral estimation. This intermediate step is similar to previous works [17, 22, 54].

Proposition 3.1. *Let the kernel* $\varphi_{\mathbf{r}}(t)=0, t\notin\mathcal{D}\subset[0,\tau)$, *be known, and further, let us assume that* $\varphi_{\mathbf{r}}(t)$ *satisfies Strang–Fix properties,* $e^{-\frac{\nu_m t}{T}}=\sum_{n\in\mathbb{Z}}\mu_{m,n}\overline{\varphi}_{\mathbf{r}}(t-nT)$, $m\in\mathbb{I}_M$. *Given the measurements defined as* $g_{\mathbf{r}}[n]=(s_{\mathbf{r}}*\varphi_{\mathbf{r}})(nT)\equiv\sum_{k=0}^{K-1}\Gamma_{\mathbf{r}}[k]\varphi_{\mathbf{r}}(nT-\tau_{\mathbf{r}}[k])$, *then the SRF* $s_{\mathbf{r}}(t)=\sum_{k=0}^{K-1}\Gamma_{\mathbf{r}}[k]\delta(t-\tau_{\mathbf{r}}[k])$ *can be recovered from* $N\geqslant 2K$ *samples.*

Proof. As detailed in (3.9), Lemma 3.2, the exponential reproducing coefficients $\{\mu_{m,n}\}_{m\in\mathbb{I}_M}^{n\in\mathbb{I}_N}$ can be computed given $\varphi_{\mathbf{r}}$. With coefficients $\{\mu_{m,n}\}_{m\in\mathbb{I}_M}^{n\in\mathbb{I}_N}$ known, let us define exponential moments,

$$\begin{aligned} y_{\mathbf{r}}[m] &= \sum_{n=0}^{N-1}\mu_{m,n}g_{\mathbf{r}}[n] \overset{(2.3)}{=} \sum_{n=0}^{N-1}\mu_{m,n}\sum_{k=0}^{K-1}\Gamma_{\mathbf{r}}[k]\,\varphi_{\mathbf{r}}(nT-\tau_{\mathbf{r}}[k]) \\ &= \sum_{k=0}^{K-1}\Gamma_{\mathbf{r}}[k]\sum_{n=0}^{N-1}\mu_{m,n}\overline{\varphi}_{\mathbf{r}}(\tau_{\mathbf{r}}[k]-nT) \overset{(3.1)}{=} \sum_{k=0}^{K-1}\Gamma_{\mathbf{r}}[k]e^{\frac{-\nu_m}{T}\tau_{\mathbf{r}}[k]}, \end{aligned} \tag{3.3}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

which is a finite sum of K-complex exponentials. Let $\nu_m = \jmath\frac{2m\pi}{N}, m \in \mathbb{I}_M$. The *unknown* frequencies in (3.3) can be found by using Prony's method as follows. Let $h_{\mathbf{r}}[m]$ be the filter with z-transform,

$$\mathsf{H}_{\mathbf{r}}(z) = \sum_{m=0}^{K} h[m]\, z^{-m} = \prod_{k=0}^{K-1} (1 - u_k z^{-1}),\ u_k = e^{\frac{-\jmath 2\pi}{\tau}\tau_{\mathbf{r}}[k]} \text{ and } \mathsf{H}_{\mathbf{r}} \in P_K, \tag{3.4}$$

where $\{u_k\}_{k\in\mathbb{I}_K}$ are the roots of $\mathsf{H}_{\mathbf{r}}$ since $\mathsf{H}_{\mathbf{r}}(u_k) = 0, k \in \mathbb{I}_K$. Then, $h[m]$ annihilates $y_{\mathbf{r}}[m]$ or

$$(h * y_{\mathbf{r}})[m] = \sum_{l=0}^{K} h[l] y_{\mathbf{r}}[m-l] = \sum_{k=0}^{K-1} \Gamma_{\mathbf{r}}[k] \left(\sum_{l=0}^{K} h[l] u_k^{-l} \right) u_k^m = 0. \tag{3.5}$$

The annihilation filter $\{h[m]\}_{m\in\mathbb{I}_{K+1}}$ can be found by solving a system of linear equations in (3.5) while $\{\tau_{\mathbf{r}}[k]\}_{k\in\mathbb{I}_K}$ can be obtained by computing the zeros u_k of the polynomial $\mathsf{H}_{\mathbf{r}}$ constructed in (3.4). The amplitudes $\{\Gamma_{\mathbf{r}}[k]\}_{k\in\mathbb{I}_K}$ can be computed via least-squares since both $\{y_{\mathbf{r}}[m]\}_{m\in\mathbb{I}_M}$ and $\{\tau_{\mathbf{r}}[k]\}_{k\in\mathbb{I}_K}$ are already known. The problem can be solved as soon as there are at least as many equations as unknowns, i.e., $N \geqslant 2K$ samples for estimating $2K$ unknowns $\{\Gamma_{\mathbf{r}}[k], \tau_{\mathbf{r}}[k]\}_{k\in\mathbb{I}_K}$. ■

Until this point, we have seen that kernels that satisfy Strang–Fix properties can lead to sparse SRes recovery. That said, it still remains to justify why ToF kernels can reproduce exponentials and what is a numerical method to compute the exponential reproducing coefficients $\{\mu_{m,n}\}_{m\in\mathbb{I}_M}^{n\in\mathbb{I}_N}$.

ToF kernels satisfy Strang–Fix conditions. In this part, we demonstrate that the common features of the ToF kernel, i.e., *time-localization* and *smoothness,* mathematically result in a decaying Fourier spectrum, and thus the Strang–Fix conditions in (3.2) hold approximately. More specifically, we have the following:

- Time-localization: Since the kernel satisfies $\varphi_{\mathbf{r}}(t) = 0,\ t \notin \mathcal{D} \subset \mathbb{R}$, we can express $\varphi_{\mathbf{r}}(t)$ as

$$\varphi_{\mathbf{r}}(t) = \sum_{m\in\mathbb{Z}} \breve{\varphi}_{\mathbf{r}}[m] e^{\jmath\frac{2m\pi t}{\tau}},\ \breve{\varphi}_{\mathbf{r}}[m] = \frac{1}{\tau}\int_0^{\tau} \varphi_{\mathbf{r}}(t) e^{-\jmath\frac{2m\pi t}{\tau}} dt, \text{ and } \tau > \max_{k,\mathbf{r}} |\tau_{\mathbf{r}}[k]|. \tag{3.6}$$

- Smoothness: Let $\varphi_{\mathbf{r}} \in C(\mathbb{R})$ be a time-localized kernel with $\varphi_{\mathbf{r}}(0) = \varphi_{\mathbf{r}}(\tau)$;[5] then its Fourier series coefficients decay as frequency increases, i.e.,

$$|\breve{\varphi}_{\mathbf{r}}[m]| \leqslant \frac{\tau}{|2m\pi|} \left\| \partial_t^{(1)} \varphi_{\mathbf{r}}(t) \right\|_{\infty},\ m \in \mathbb{Z}/\{0\}. \tag{3.7}$$

A consequence of these two properties is that $\{\breve{\varphi}_{\mathbf{r}}[m]\}_{m\in\mathbb{Z}}$ contract to zero as $|m|$ increases, i.e.,

$$\forall \epsilon > 0, \exists M \in \mathbb{N}^{+} \Rightarrow |\breve{\varphi}_{\mathbf{r}}[m]| < \epsilon, |m| > M \tag{3.8}$$

---

[5] Without loss of any generality, we use the boundary condition to simplify the result in (3.7), which is a standard result in Fourier analysis.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

and hence $\varphi_{\mathbf{r}}$ is well approximated by a finite number of Fourier series coefficients. That is to say, (3.8) implicitly mimics Strang–Fix conditions in (3.2) ($H=0$) since for every $|m|>M$, we have $|\breve{\varphi}_{\mathbf{r}}[m]|\approxeq 0 \implies |\mathcal{L}_{\varphi_{\mathbf{r}}}(\jmath\frac{2m\pi}{\tau})|\approxeq 0$, with an error bounded by $\frac{\tau}{|2M\pi|}||\partial_t^{(1)}\varphi_{\mathbf{r}}(t)||_\infty$. Hence, $\varphi_{\mathbf{r}}$ approximately satisfies the Strang–Fix conditions in (3.2).

Computation of exponential reproducing coefficients. Given the kernel $\varphi_{\mathbf{r}}(t)=(p_{\mathbf{r}}*\psi_{\mathbf{r}})(t), t\in[0,\tau)$, we now present an efficient method to compute exponential reproducing coefficients.

Lemma 3.2. *Let the kernel* $\varphi_{\mathbf{r}}(t)=0, t\notin\mathcal{D}\subset[0,\tau)$, *be known, and further, let us assume that* $\varphi_{\mathbf{r}}(t)$ *satisfies Strang–Fix properties,* $e^{-\frac{\nu_m t}{T}}=\sum_{n\in\mathbb{I}_N}\mu_{m,n}\overline{\varphi}_{\mathbf{r}}(t-nT)$, $m\in\mathbb{I}_M$. *Let* $\nu_m=\jmath\frac{2m\pi}{N}, m\in\mathbb{I}_M$. *Then, the coefficients* $\{\mu_{m,n}\}_{n\in\mathbb{I}_N}^{m\in\mathbb{I}_M}$ *are given by*

$$\mathbf{U}=\frac{1}{N}\mathscr{D}_K{}^{-1}(\breve{\boldsymbol{\varphi}}_{\mathbf{r}})\,\mathbf{W}_M^N,\ \mathbf{U}=[\mu_{m,n}]_{m\in\mathbb{I}_M}^{n\in\mathbb{I}_N},\ \textit{and}\ \tau=NT. \tag{3.9}$$

Proof. From the exponential reproduction properties in (3.1), we obtain that

$$\int_0^\tau\sum_{n=0}^{N-1}\mu_{m,n}\overline{\varphi}_{\mathbf{r}}(t-nT)e^{\frac{\nu_l t}{T}}\,dt\overset{(3.1)}{=}\int_0^{NT}e^{(\nu_l-\nu_m)\frac{t}{T}}\,dt=\tau\delta\left[l-m\right], \tag{3.10}$$

where $\delta[\cdot]$ is the Kronecker delta sequence. Since $\nu_m=\jmath\frac{2m\pi}{N}, m\in\mathbb{I}_M$, then $\varphi_{\mathbf{r}}$ satisfies that

$$\int_0^\tau\overline{\varphi}_{\mathbf{r}}\left(t-nT\right)e^{\frac{\nu_l t}{T}}\,dt=\tau\breve{\varphi}_{\mathbf{r}}\left[l\right]e^{\nu_l n}. \tag{3.11}$$

Hence, combining (3.10) and (3.11), we have that

$$\int_0^\tau\sum_{n=0}^{N-1}\mu_{m,n}\overline{\varphi}_{\mathbf{r}}(t-nT)e^{\frac{\nu_l t}{T}}\,dt=\tau\sum_{n=0}^{N-1}\mu_{m,n}\breve{\varphi}_{\mathbf{r}}\left[l\right]e^{\nu_l n}=\tau\delta\left[l-m\right],\ l,m\in\mathbb{I}_M, \tag{3.12}$$

which can be algebraically rewritten in matrix form as $\mathbf{U}\mathbf{V}_N^M\mathscr{D}_K(\breve{\boldsymbol{\varphi}}_{\mathbf{r}})=\frac{1}{N}\mathbf{I}$, where $\mathbf{I}\in\mathbb{R}^{M\times M}$ is the identity matrix. By simplification, we eventually have that

$$\underset{(M\times N)}{\boxed{\quad\mathbf{U}\quad}}=\frac{1}{N}\underset{(M\times M)}{\boxed{\mathscr{D}_K^{-1}(\breve{\boldsymbol{\varphi}}_{\mathbf{r}})}}\underset{(M\times N)}{\boxed{\quad\mathbf{W}_M^N\quad}}$$

$$\mathbf{U}=\frac{1}{N}\mathscr{D}_K{}^{-1}(\breve{\boldsymbol{\varphi}}_{\mathbf{r}})\,\mathbf{W}_M^N\quad\text{where}\quad\left[\mathscr{D}_K\left(\breve{\boldsymbol{\varphi}}_{\mathbf{r}}\right)\right]_{m,m}=\breve{\varphi}_{\mathbf{r}}\left[m\right]. \tag{3.13}$$ ■

4. Blind super-resolved ToF imaging.

In the intermediate scenario of a known kernel, we demonstrate that the SRes ToF imaging can be achieved by employing the exponential reproduction properties of the kernel. In this section, we investigate the blind ToF imaging methodology that allows for adaptive estimation of both the SRF and the kernel. In the absence of kernel calibration, our setup takes the shape of a BSD problem, which is generally challenging to solve due to its ill-posed formulation. In order to tackle this problem, we opt for an alternating minimization framework that splits the BSD into two tractable subproblems, as described in Figure 5.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

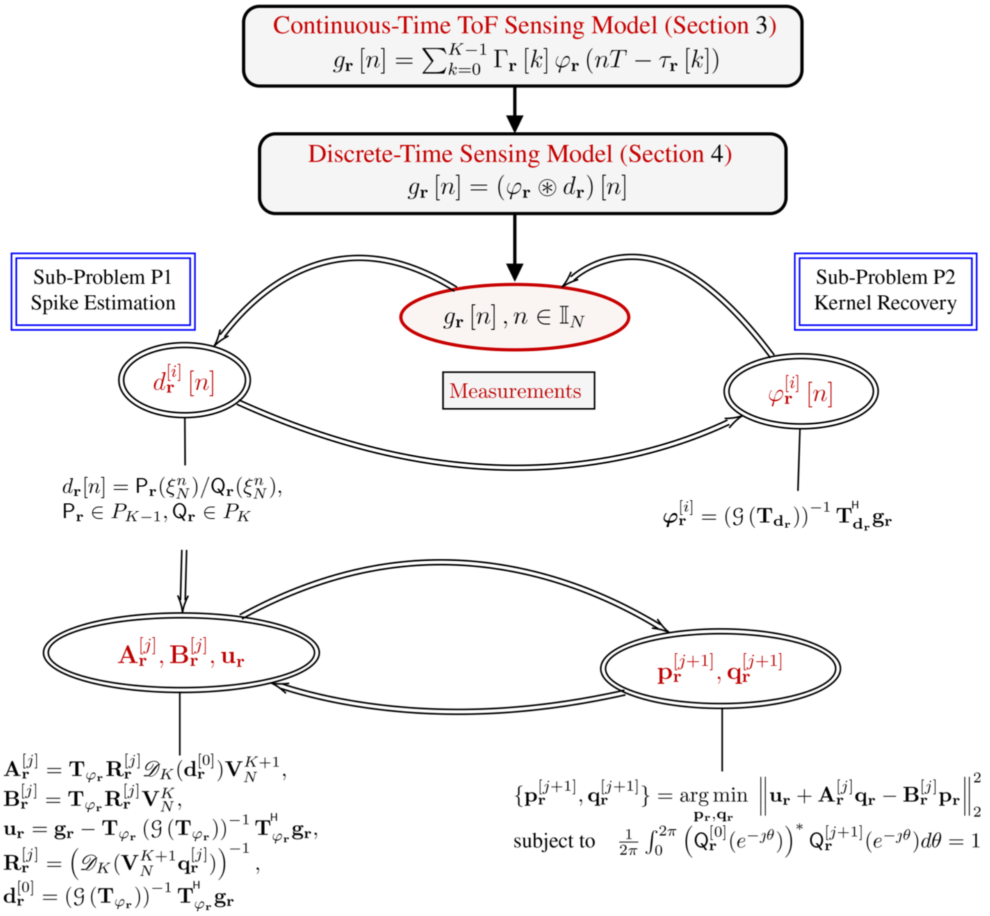


Figure 5. *Block diagram of the proposed blind ToF imaging methodology. In the absence of kernel calibration, our setting takes the form of a BSD problem, where we split it into two tractable subproblems:* $\boxed{\text{P1}}$ *that addresses recovery of* $d_{\mathbf{r}}[n]$ *via continuous-time spike estimation (see section 4.2) and* $\boxed{\text{P2}}$ *that solves for* $\varphi_{\mathbf{r}}[n]$ *utilizing least-squares fitting (see section 4.3).*

4.1. Sparse model for ToF measurements.

We first focus on a special subset of Strang–Fix kernels, i.e., bandlimited kernels. This scenario reveals an exact model that relates discrete ToF measurements $\{g_{\mathbf{r}}[n]\}_{n\in\mathbb{I}_N}$ to a continuous-time spike representation. When the bandlimitedness constraint is relaxed, we are left with an approximate model which, as shown via experiments, serves as a compelling match to real-world scenarios.

Proposition 4.1. *Let the kernel* $\varphi_{\mathbf{r}}$ *in* (3.6) *be bandlimited, i.e.,* $\breve{\varphi}_{\mathbf{r}}[m] = 0, |m| > L$. *Then,*

$$\begin{aligned} g_{\mathbf{r}}[n] &= (s_{\mathbf{r}} * \varphi_{\mathbf{r}})(nT) \equiv \sum_{k=0}^{K-1} \Gamma_{\mathbf{r}}[k] \varphi_{\mathbf{r}}(nT - \tau_{\mathbf{r}}[k]) \\ &\equiv (\varphi_{\mathbf{r}} \circledast d_{\mathbf{r}})[n] \quad \text{where} \quad d_{\mathbf{r}}[n] \overset{\text{def}}{=} \frac{1}{N} \sum_{k=0}^{K-1} \frac{\Gamma_{\mathbf{r}}[k](1-(u_k)^N)}{(1-u_k\xi_N^n)} = \frac{\mathsf{P}_{\mathbf{r}}(\xi_N^n)}{\mathsf{Q}_{\mathbf{r}}(\xi_N^n)} \end{aligned} \tag{4.1}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

and where $\circledast$ *denotes circular convolution,* $\mathsf{P}_{\mathbf{r}} \in P_{K-1}, \mathsf{Q}_{\mathbf{r}} \in P_K$, $u_k \overset{(3.4)}{=} e^{\frac{-\jmath 2\pi}{\tau}\tau_{\mathbf{r}}[k]}$, *and* $\xi_N^n = e^{\jmath\frac{2\pi n}{N}}$.

Proof. Our starting point is the low-pass structure of the kernel $\varphi_{\mathbf{r}}$, resulting in

$$\varphi_{\mathbf{r}}[n] = \varphi_{\mathbf{r}}(t)|_{t=nT} = \langle \varphi_{\mathbf{r}}(t), f_L(t-nT)\rangle, n \in \mathbb{I}_N, \tag{4.2}$$

where $f_L(t) = \sin((2L+1)\pi t/\tau)/\sin(\pi t/\tau)$ is the low-pass filter.

Let $\widehat{\varphi}_{\mathbf{r}}$ denote the DFT of $\varphi_{\mathbf{r}}$. We show that the DFT coefficients $\{\widehat{\varphi}_{\mathbf{r}}[l]\}_{l\in\mathbb{I}_M}$ and their continuous counterpart $\{\breve{\varphi}_{\mathbf{r}}[l]\}_{l\in\mathbb{I}_M}$ are linearly dependent, since

$$\begin{aligned}\widehat{\varphi}_{\mathbf{r}}[l] &= \sum_{n=0}^{N-1}\varphi_{\mathbf{r}}[n]\xi_N^{-n\cdot l} \overset{(4.2)}{=} \left\langle \varphi_{\mathbf{r}}(t), \sum_{n=0}^{N-1}\sum_{|m|\leqslant L} e^{-\jmath\frac{2m\pi t}{\tau}} e^{\jmath\frac{2n\pi(m-l)}{N}}\right\rangle \\ &= \left\langle \varphi_{\mathbf{r}}(t), \sum_{|m|\leqslant L} e^{-\jmath\frac{2m\pi t}{\tau}}\delta[l-m]\right\rangle = \left\langle \varphi_{\mathbf{r}}(t), e^{-\jmath\frac{2l\pi t}{\tau}}\right\rangle \overset{(3.6)}{=} \tau\breve{\varphi}_{\mathbf{r}}[l].\end{aligned} \tag{4.3}$$

Similarly, we have $\widehat{g}_{\mathbf{r}}[l] = \tau\breve{g}_{\mathbf{r}}[l]$ where $\breve{g}_{\mathbf{r}}[m] = \frac{1}{\tau}\int_0^\tau g_{\mathbf{r}}(t)e^{-\jmath\frac{2m\pi t}{\tau}}dt$. Moreover, notice that

$$\begin{aligned}g_{\mathbf{r}}(t) &= \sum_{m\in\mathbb{Z}}\breve{g}_{\mathbf{r}}[m]e^{\jmath\frac{2m\pi t}{\tau}} = \sum_{k=0}^{K-1}\Gamma_{\mathbf{r}}[k]\,\varphi_{\mathbf{r}}(t-\tau_{\mathbf{r}}[k]) \\ &= \sum_{m\in\mathbb{Z}}\left(\breve{\varphi}_{\mathbf{r}}[m]\sum_{k=0}^{K-1}\Gamma_{\mathbf{r}}[k]\,e^{-\jmath\frac{2m\pi\tau_{\mathbf{r}}[k]}{\tau}}\right)e^{\jmath\frac{2m\pi t}{\tau}}.\end{aligned}$$

Hence, $\breve{g}_{\mathbf{r}}[m]$ can be expressed in terms of $\breve{\varphi}_{\mathbf{r}}[m]$ and $\{\Gamma_{\mathbf{r}}[k], \tau_{\mathbf{r}}[k]\}_{k\in\mathbb{I}_K}$ as

$$\breve{g}_{\mathbf{r}}[m] = \breve{\varphi}_{\mathbf{r}}[m]\breve{s}_{\mathbf{r}}[m],\ \breve{s}_{\mathbf{r}}[m] = \sum_{k=0}^{K-1}\Gamma_{\mathbf{r}}[k]\,e^{-\jmath\frac{2m\pi\tau_{\mathbf{r}}[k]}{\tau}},\quad m\in\mathbb{Z}. \tag{4.4}$$

Combining (4.4) and (4.3), we eventually derive the interplay between $\widehat{g}_{\mathbf{r}}$ and $\widehat{\varphi}_{\mathbf{r}}$ as

$$\widehat{g}_{\mathbf{r}}[l] = \tau\breve{g}_{\mathbf{r}}[l] \overset{(4.4)}{=} \widehat{\varphi}_{\mathbf{r}}[l]\breve{s}_{\mathbf{r}}[l] = \widehat{\varphi}_{\mathbf{r}}[l]\sum_{k=0}^{K-1}\Gamma_{\mathbf{r}}[k]\,e^{-\jmath\frac{2l\pi\tau_{\mathbf{r}}[k]}{\tau}},\ l\in\mathbb{I}_N. \tag{4.5}$$

(4.5) can be recast in the time domain, which gives rise to a sparse representation

$$\begin{aligned}g_{\mathbf{r}}[n] &= \frac{1}{N}\sum_{l=0}^{N-1}\widehat{g}_{\mathbf{r}}[l]\xi_N^{n\cdot l} = \frac{1}{N}\sum_{l=0}^{N-1}\sum_{u=0}^{N-1}\varphi_{\mathbf{r}}[u]\xi_N^{-u\cdot l}\breve{s}_{\mathbf{r}}[l]\xi_N^{n\cdot l} \\ &= \sum_{u=0}^{N-1}\varphi_{\mathbf{r}}[u]\left(\frac{1}{N}\sum_{l=0}^{N-1}\breve{s}_{\mathbf{r}}[l]\xi_N^{(n-u)\cdot l}\right) = \sum_{u=0}^{N-1}\varphi_{\mathbf{r}}[u]d_{\mathbf{r}}[(n-u)_{\text{mod } N}],\end{aligned} \tag{4.6}$$

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

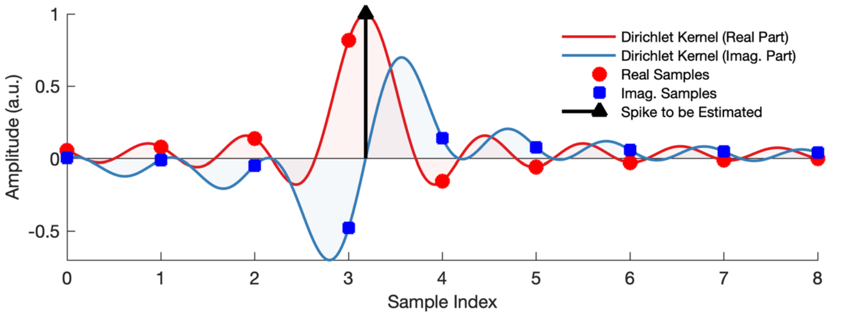


Figure 6. *Visualization of* $d_{\mathbf{r}}$ *in* (4.1).

where mod denotes modulo operation and $d_{\mathbf{r}}$ is parametrized by $\{\tau_{\mathbf{r}}[k], \Gamma_{\mathbf{r}}[k]\}_{k\in\mathbb{I}_K}$ (see Figure 6) as

$$
\begin{aligned}
d_{\mathbf{r}}[n] &= \frac{1}{N}\sum_{l=0}^{N-1} \breve{s}_{\mathbf{r}}[l]\xi_N^{n\cdot l} \overset{(4.4)}{=} \frac{1}{N}\sum_{k=0}^{K-1}\Gamma_{\mathbf{r}}[k]\sum_{l=0}^{N-1} e^{\jmath\frac{2l\pi}{N}\left(n-\frac{\tau_{\mathbf{r}}[k]}{T}\right)} \\
&= \sum_{k=0}^{K-1}\frac{\Gamma_{\mathbf{r}}[k](1-(u_k)^N)}{N(1-u_k\xi_N^n)} = \frac{\mathsf{P}_{\mathbf{r}}(\xi_N^n)}{\mathsf{Q}_{\mathbf{r}}(\xi_N^n)}, \quad \mathsf{P}_{\mathbf{r}} \in P_{K-1}, \mathsf{Q}_{\mathbf{r}} \in P_K.
\end{aligned}
\tag{4.7}
$$

In the above, $\mathsf{Q}_{\mathbf{r}}$ is related to $\mathsf{H}_{\mathbf{r}}$ in (3.4) as $\mathsf{Q}_{\mathbf{r}}(z^{-1}) = \mathsf{H}_{\mathbf{r}}(z) \Longleftrightarrow \mathsf{Q}_{\mathbf{r}}(u_k^{-1}) = 0, k \in \mathbb{I}_K$. ■

When considering the ToF kernels discussed in section 3, the model in (4.1) thus becomes an accurate *approximation* due to the decaying Fourier spectrum of the kernels, as shown in (3.7). This enables a flexible and practical framework that leverages the common kernel characteristics and recovers sparse spikes on a continuum. More specifically, on the one hand, this derived model matches the various ToF modalities, which we validate via extensive hardware experiments in section 5. On the other hand, this model leads to a tractable and efficient optimization method for the SRF and kernel recovery. Consequently, in the presence of distortions, e.g., quantization resolution, system noise, and interreflections, the blind ToF imaging problem can be posed as

$$
\min_{\varphi_{\mathbf{r}}, \mathsf{P}_{\mathbf{r}}, \mathsf{Q}_{\mathbf{r}}} \|\mathbf{g}_{\mathbf{r}} - \varphi_{\mathbf{r}} \circledast \mathbf{d}_{\mathbf{r}}\|_2^2, \text{ s.t. } d_{\mathbf{r}}[n] = \frac{\mathsf{P}_{\mathbf{r}}(\xi_N^n)}{\mathsf{Q}_{\mathbf{r}}(\xi_N^n)}, \ \mathsf{P}_{\mathbf{r}} \in P_{K-1},\ \mathsf{Q}_{\mathbf{r}} \in P_K,\ n \in \mathbb{I}_N. \tag{4.8}
$$

What makes the minimization in (4.8) nontrivial are the structure and the constraints of the setup. In order to address this problem, we employ an alternating minimization strategy where the goal is to split (4.8) into two tractable subproblems: $\boxed{\mathsf{P1}}$ that addresses the recovery of $\mathsf{P}_{\mathbf{r}}, \mathsf{Q}_{\mathbf{r}}$ via nonlinear model-fitting and $\boxed{\mathsf{P2}}$ that solves for $\varphi_{\mathbf{r}}$ via least-squares fitting.

4.2. Subproblem P1: Continuous-time spike estimation.

Assuming that $\varphi_{\mathbf{r}}$ is known, it remains to estimate $\mathsf{P}_{\mathbf{r}}, \mathsf{Q}_{\mathbf{r}}$ by solving the following quadratic minimization:

$$
\boxed{\mathsf{P1}} \qquad \min_{\mathsf{P}_{\mathbf{r}}, \mathsf{Q}_{\mathbf{r}}} \|\mathbf{g}_{\mathbf{r}} - \varphi_{\mathbf{r}} \circledast \mathbf{d}_{\mathbf{r}}\|_2^2, \text{ s.t. } d_{\mathbf{r}}[n] = \frac{\mathsf{P}_{\mathbf{r}}(\xi_N^n)}{\mathsf{Q}_{\mathbf{r}}(\xi_N^n)}, \ \mathsf{P}_{\mathbf{r}} \in P_{K-1},\ \mathsf{Q}_{\mathbf{r}} \in P_K,\ n \in \mathbb{I}_N. \tag{4.9}
$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

The bilinearity of convolution operation results in

$$\varphi_{\mathbf{r}} \circledast \mathbf{d}_{\mathbf{r}} = \mathbf{T}_{\varphi_{\mathbf{r}}} \mathbf{d}_{\mathbf{r}}, \;\; [\mathbf{T}_{\varphi_{\mathbf{r}}}]_{m,n} = \varphi_{\mathbf{r}}[(m-n)_{\operatorname{mod} N}].$$

Then, $\boxed{\mathsf{P1}}$ in (4.9) translates to

$$\min_{\mathsf{P}_{\mathbf{r}},\mathsf{Q}_{\mathbf{r}}} \sum_{m=0}^{N-1} \left| g_{\mathbf{r}}[m] - \sum_{n=0}^{N-1} [\mathbf{T}_{\varphi_{\mathbf{r}}}]_{m,n} \frac{\mathsf{P}_{\mathbf{r}}(\xi_N^n)}{\mathsf{Q}_{\mathbf{r}}(\xi_N^n)} \right|^2 . \tag{4.10}$$

As solving (4.10) is challenging due to its nonlinearity, we opt for an iterative strategy by constructing a collection of estimates for $\mathsf{P}_{\mathbf{r}}, \mathsf{Q}_{\mathbf{r}}$ and selecting the one that minimizes the MSE of the measurements via (4.10). These estimates $\{\mathsf{P}_{\mathbf{r}}^{[j+1]}, \mathsf{Q}_{\mathbf{r}}^{[j+1]}\}$ are found iteratively by solving the following approximate problem (since $\mathsf{Q}_{\mathbf{r}}^{[j]} \approx \mathsf{Q}_{\mathbf{r}}$):

$$\begin{gathered} \min_{\mathsf{P}_{\mathbf{r}},\mathsf{Q}_{\mathbf{r}}} \sum_{m=0}^{N-1} \left| u_{\mathbf{r}}[m] + \sum_{n=0}^{N-1} [\mathbf{T}_{\varphi_{\mathbf{r}}}]_{m,n} \frac{d_{\mathbf{r}}^{[0]}[n]\mathsf{Q}_{\mathbf{r}}(\xi_N^n) - \mathsf{P}_{\mathbf{r}}(\xi_N^n)}{\mathsf{Q}_{\mathbf{r}}^{[j]}(\xi_N^n)} \right|^2 , \\ \text{where } u_{\mathbf{r}}[m] = g_{\mathbf{r}}[m] - \sum_{n=0}^{N-1} [\mathbf{T}_{\varphi_{\mathbf{r}}}]_{m,n} \, d_{\mathbf{r}}^{[0]}[n], \\ \mathbf{d}_{\mathbf{r}}^{[0]} = \left(\mathcal{G}\left(\mathbf{T}_{\varphi_{\mathbf{r}}}\right)\right)^{-1} \mathbf{T}_{\varphi_{\mathbf{r}}}^{\mathsf{H}} \mathbf{g}_{\mathbf{r}}, \quad j \in \mathbb{I}_{j_{\max}}, \end{gathered} \tag{4.11}$$

where $\mathbf{d}_{\mathbf{r}}^{[0]}$ is an initial estimate of $\mathbf{d}_{\mathbf{r}}$ and the resulting estimation error is characterized by the residue entry $\mathbf{u}_{\mathbf{r}}$. This structure ensures that the solution to (4.11) minimizes the model-fitting error defined in (4.9). Different initialization of $\mathsf{Q}_{\mathbf{r}}^{[0]}$ results in diverse estimates of $\mathsf{Q}_{\mathbf{r}}^{[j]}$, improving the accuracy and robustness of the algorithm. Here, we provide a deterministic initialization strategy: we pick the K most prominent peaks of $\mathbf{d}_{\mathbf{r}}^{[0]}$ [30], which provides the roots of $\mathsf{Q}_{\mathbf{r}}^{[j]}$ as well as the polynomial coefficients. It is possible that the stopping criterion in (4.20) may not be met for this choice of $\mathsf{Q}_{\mathbf{r}}^{[0]}$ after reaching the maximum iteration count $j_{\max}$. In such cases, the algorithm is restarted with random initialized $\mathsf{Q}_{\mathbf{r}}^{[0]}$ that follows independent and identical unit Gaussian distributions.

Having estimated $\mathsf{Q}_{\mathbf{r}}$ and $\mathsf{P}_{\mathbf{r}}$, spike locations $\tau_{\mathbf{r}}[k]$ are obtained by $\mathsf{roots}(\mathsf{Q}_{\mathbf{r}}) \mapsto u_k^{-1}$ since $\mathsf{Q}_{\mathbf{r}}(u_k^{-1}) = 0$. The corresponding amplitudes $\Gamma_{\mathbf{r}}[k]$ are obtained via least-squares,

$$\tau_{\mathbf{r}}[k] = -\frac{\tau}{2\pi} \mathsf{Im}(\log(u_k)) \quad \text{and} \quad \Gamma_{\mathbf{r}}[k] = -\frac{N u_k \mathsf{P}_{\mathbf{r}}\left(u_k^{-1}\right)}{\left(1 - u_k^N\right) \partial_z^{(1)} \mathsf{Q}_{\mathbf{r}}\left(z\right)\Big|_{z=u_k^{-1}}}. \tag{4.12}$$

With $\{\tau_{\mathbf{r}}[k], \Gamma_{\mathbf{r}}[k]\}_{k \in \mathbb{I}_K}$, $\mathbf{d}_{\mathbf{r}}$ can be reconstructed using (4.1) and 3D SRF can be recovered via (2.1).

Algorithmic implementation. We provide an algorithm for (4.11). The trigonometric polynomials $\mathsf{P}_{\mathbf{r}}(\xi_N^n), \mathsf{Q}_{\mathbf{r}}(\xi_N^n)$ in (4.11) can be algebraically written as

$$[\mathsf{P}_{\mathbf{r}}(\xi_N^n)] = \mathbf{V}_N^K \mathbf{p}_{\mathbf{r}} \;\; \text{and} \;\; [\mathsf{Q}_{\mathbf{r}}(\xi_N^n)] = \mathbf{V}_N^{K+1} \mathbf{q}_{\mathbf{r}},$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

where $\mathbf{p_r}$ and $\mathbf{q_r}$ are the coefficients of $\mathsf{P}_\mathbf{r} \in P_{K-1}$ and $\mathsf{Q}_\mathbf{r} \in P_K$. Assuming the estimate $\mathbf{q}_\mathbf{r}^{[j]}$ of $\mathbf{q_r}$ at iteration-j is known, the minimization at iteration-$j+1$ can be reformulated in matrix form as

$$\{\mathbf{p}_\mathbf{r}^{[j+1]}, \mathbf{q}_\mathbf{r}^{[j+1]}\} = \arg\min_{\mathbf{p_r},\mathbf{q_r}} \left\| \mathbf{u_r} + \mathbf{A}_\mathbf{r}^{[j]} \mathbf{q_r} - \mathbf{B}_\mathbf{r}^{[j]} \mathbf{p_r} \right\|_2^2, \tag{4.13}$$

where the entities $\mathbf{A}_\mathbf{r}^{[j]}$, $\mathbf{B}_\mathbf{r}^{[j]}$, and $\mathbf{u_r}$ are respectively given by

(4.14)

$$\begin{aligned} &\mathbf{A}_\mathbf{r}^{[j]} = \mathbf{T}_{\varphi_\mathbf{r}} \mathbf{R}_\mathbf{r}^{[j]} \mathscr{D}_K(\mathbf{d}_\mathbf{r}^{[0]}) \mathbf{V}_N^{K+1}, \qquad \mathbf{B}_\mathbf{r}^{[j]} = \mathbf{T}_{\varphi_\mathbf{r}} \mathbf{R}_\mathbf{r}^{[j]} \mathbf{V}_N^{K}, \quad \mathbf{u_r} = \mathbf{g_r} - \mathbf{T}_{\varphi_\mathbf{r}} \left(\mathcal{G}\left(\mathbf{T}_{\varphi_\mathbf{r}}\right)\right)^{-1} \mathbf{T}_{\varphi_\mathbf{r}}^{\mathsf{H}} \mathbf{g_r}, \\ &\text{and in the above, the variables} \quad \mathbf{R}_\mathbf{r}^{[j]} = \left(\mathscr{D}_K(\mathbf{V}_N^{K+1} \mathbf{q}_\mathbf{r}^{[j]})\right)^{-1}, \text{ and } \mathbf{d}_\mathbf{r}^{[0]} = \left(\mathcal{G}\left(\mathbf{T}_{\varphi_\mathbf{r}}\right)\right)^{-1} \mathbf{T}_{\varphi_\mathbf{r}}^{\mathsf{H}} \mathbf{g_r}. \end{aligned}$$

We adopt a normalization constraint below to ensure the uniqueness of the optimal solution to (4.13),

$$\frac{1}{2\pi} \int_0^{2\pi} \left(\mathsf{Q}_\mathbf{r}^{[0]}(e^{-\jmath\theta})\right)^* \mathsf{Q}_\mathbf{r}^{[j+1]}(e^{-\jmath\theta}) d\theta = 1, \tag{4.15}$$

where $(\cdot)^*$ denotes the conjugation operator and $\mathsf{Q}_\mathbf{r}^{[0]}$ is the initialization of the algorithm. Consequently, the quadratic minimization (4.13) can be posed as

$$\begin{aligned} \{\mathbf{p}_\mathbf{r}^{[j+1]}, \mathbf{q}_\mathbf{r}^{[j+1]}\} = \arg\min_{\mathbf{p_r},\mathbf{q_r}} & \left\| \mathbf{u_r} + \mathbf{A}_\mathbf{r}^{[j]} \mathbf{q_r} - \mathbf{B}_\mathbf{r}^{[j]} \mathbf{p_r} \right\|_2^2 \\ \text{subject to } & \frac{1}{2\pi} \int_0^{2\pi} \left(\mathsf{Q}_\mathbf{r}^{[0]}(e^{-\jmath\theta})\right)^* \mathsf{Q}_\mathbf{r}^{[j+1]}(e^{-\jmath\theta}) d\theta = 1 \end{aligned} \tag{4.16}$$

for which its optimal solution can be found by solving the following system of linear equations:

$$\begin{bmatrix} \mathcal{G}\left(\mathbf{C}_\mathbf{r}^{[j]}\right) & \mathbf{X}_\mathbf{r}^{[0]} \\ (\mathbf{X}_\mathbf{r}^{[0]})^{\mathsf{H}} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{X}_\mathbf{r}^{[j+1]} \\ \lambda \end{bmatrix} = \begin{bmatrix} (\mathbf{C}_\mathbf{r}^{[j]})^{\mathsf{H}} \mathbf{u_r} \\ 1 \end{bmatrix} \text{ and } \quad \begin{aligned} &\mathbf{X}_\mathbf{r}^{[j+1]} = \begin{bmatrix} \mathbf{q}_\mathbf{r}^{[j+1]} \\ \mathbf{p}_\mathbf{r}^{[j+1]} \end{bmatrix}, \ \mathbf{X}_\mathbf{r}^{[0]} = \begin{bmatrix} \mathbf{q}_\mathbf{r}^{[0]} \\ \mathbf{0} \end{bmatrix}, \\ &\mathbf{C}_\mathbf{r}^{[j]} = [-\mathbf{A}_\mathbf{r}^{[j]}, \mathbf{B}_\mathbf{r}^{[j]}], \end{aligned} \tag{4.17}$$

where $\mathbf{q}_\mathbf{r}^{[0]}$ are the coefficients of $\mathsf{Q}_\mathbf{r}^{[0]} \in P_K$ and λ is the Lagrange multiplier such that the normalization constraint (4.15) is satisfied.[6]

4.3. Subproblem P2: Kernel recovery. With $\mathbf{d_r}$ known from the method in section 4.2, the minimization on $\boldsymbol{\varphi}_\mathbf{r}$ essentially boils down to the least-squares problem, resulting in the solution

$$\boldsymbol{\varphi}_\mathbf{r} = \left(\mathcal{G}\left(\mathbf{T}_{\mathbf{d_r}}\right)\right)^{-1} \mathbf{T}_{\mathbf{d_r}}^{\mathsf{H}} \mathbf{g_r}. \tag{4.18}$$

---

[6] For each initialization $\mathbf{q}_\mathbf{r}^{[0]}$, we update the trigonometric polynomial coefficients $\mathbf{q}_\mathbf{r}^{[j+1]}$ until the stopping criterion (4.20) is met. If (4.20) is not met for certain choices of $\mathbf{q}_\mathbf{r}^{[j+1]}$ after reaching the maximum iteration count $j_{\max}$, we restart the algorithm with a different initialization.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Stopping criterion. We initialize the proposed method in Algorithm 4.1 by computing

$$\varphi_{\mathbf{r}}^{[0]} = \mathrm{Re}\left(\mathbf{V}_N^N(\mathscr{D}_K(e^{\jmath\mathbf{w}})\mathbf{W}_N^N\mathbf{g}_{\mathbf{r}})\right), \qquad \mathbf{w} \sim \mathcal{N}(\mathbf{0},\mathbf{I}), \tag{4.19}$$

which is a reasonable initialization based on kernel features. Using (4.13), we then estimate $\{\tau_{\mathbf{r}}[k], \Gamma_{\mathbf{r}}[k]\}_{k\in\mathbb{I}_K}$ via (4.12) based on which we refine $\boldsymbol{\varphi}_{\mathbf{r}}$ via (4.18). We use

$$\|\mathbf{g}_{\mathbf{r}} - \boldsymbol{\varphi}_{\mathbf{r}} \circledast \mathbf{d}_{\mathbf{r}}\|_2 \leqslant \sigma \tag{4.20}$$

as our stopping criterion, where σ represents the data distortion level. In other words, we can only recover the signal up to a tolerance level of σ. Iterating the method offers robust estimates with SRes capability, which is validated via hardware experiments. Empirically, 10 random initializations and $j_{\max} = 20$ are sufficient to obtain an accurate solution that satisfies the stopping criterion in (4.20), provided that the spikes $\{\tau_{\mathbf{r}}[k]\}_{k\in\mathbb{I}_K}$ are relatively well-separated. In the challenging scenarios of resolving closely located spikes, more restarts with different initializations (≈ 20) are required to find a reasonable solution that fits the ToF data within a certain tolerance level σ. We would like to point out that the algorithmic complexity can be potentially improved by (1) leveraging algebraic structures of matrices $\mathbf{T}_{\varphi_{\mathbf{r}}}, \mathbf{T}_{\mathbf{d}_{\mathbf{r}}}$ and shared features of $\boldsymbol{\varphi}_{\mathbf{r}}$, (2) optimizing numerical calculation and implementation, and (3) utilizing effective dimensionality reduction techniques. We tabulate the algorithm run-time in the last column of Table 1. The algorithm is summarized in Algorithm 4.1.

While our method assumes the model order (K in (2.3)) is known, this is not strictly necessary. A practical approach leverages the underlying physics of the imaging process. Since light reflections decay according to the inverse-square law, for $K > 3$, the reflections are likely submerged in the noise floor. Thus, the algorithm can be implemented with $K = 4$, where any spurious spikes from overestimating the model order will reflect system noise. These will correspond to relatively smaller coefficients $\Gamma_{\mathbf{r}}$ in (2.3) and can be removed through hard-thresholding. Alternatively, σ in (4.20) measures the quality of fitting. Hence, one may develop a method that progressively fits the data until the stopping criterion (4.20) is met. Last, algorithms such as SORTE[7] [31] can also be combined with our approach to incorporate model-order estimation.

Table 1

Hardware based experimental parameters and performance evaluation.

Figure	Exp. no.	$\mathbf{r}$	K	T	$\Gamma_{\mathbf{r}}[k]$	$\tau_{\mathbf{r}}[k]$	$\widetilde{\Gamma}_{\mathbf{r}}[k]$	$\widetilde{\tau}_{\mathbf{r}}[k]$	MSE		PSNR $(\boldsymbol{\varphi}_{\mathbf{r}}, \widetilde{\boldsymbol{\varphi}}_{\mathbf{r}})$	Run-time
				($\times 10^{-12}$ sec.)		($\times 10^{-8}$ sec.)		($\times 10^{-8}$ sec.)			(dB)	(s)
					Known kernel		Blind		$\mathcal{E}(\boldsymbol{\Gamma}_{\mathbf{r}}, \widetilde{\boldsymbol{\Gamma}}_{\mathbf{r}})$	$\mathcal{E}(\boldsymbol{\tau}_{\mathbf{r}}, \widetilde{\boldsymbol{\tau}}_{\mathbf{r}})$		
Figure 8	I	$[60\ 60]^\top$	2	70	[1.19,0.23]	[8.45,9.44]	[1.19,0.23]	[8.44,9.44]	2.31×10^{-5}	1.87×10^{-4}	43.24	35.88
Figure 9(a)	II-(a)	$[60\ 60]^\top$	2	96.15	[0.86,0.52]	[3.03,4.19]	[0.86,0.53]	[3.02,4.19]	3.57×10^{-5}	1.16×10^{-4}	41.93	37.07
Figure 9(b)	II-(b)	$[60\ 60]^\top$	2	96.15	[0.64,0.54]	[3.23,4.20]	[0.66,0.53]	[3.21,4.20]	1.69×10^{-4}	1.40×10^{-4}	40.29	36.63
Figure 9(c)	II-(c)	$[60\ 60]^\top$	2	96.15	[0.48,0.59]	[3.42,4.19]	[0.58,0.46]	[3.43,4.18]	1.35×10^{-2}	7.13×10^{-5}	39.71	117.54
Figure 9(d)	II-(d)	$[60\ 60]^\top$	2	96.15	[0.34,0.58]	[3.64,4.20]	[0.48,0.43]	[3.62,4.20]	2.13×10^{-2}	1.90×10^{-4}	39.18	122.32
Figure 13(b)	III-(a)	$[50\ 70]^\top$	3	70	[0.70,0.44,0.13]	[8.44,9.65,10.95]	[0.66,0.48,0.16]	[8.44,9.65,10.95]	1.50×10^{-3}	1.57×10^{-5}	36.33	114.97
Figure 13(c)	III-(b)	$[45\ 60]^\top$	3	70	[0.29,0.36,0.21]	[8.51,9.71,11.08]	[0.28,0.36,0.23]	[8.53,9.74,11.08]	1.52×10^{-4}	4.70×10^{-4}	36.22	106.66
Figure 12(b)	IV-(a)	—	2	6.1	[1.34,1.72]	[1.11,1.24]	[1.33,1.72]	[1.11,1.24]	2.61×10^{-5}	2.33×10^{-8}	47.72	54.54
Figure 12(c)	IV-(b)	—	2	6.1	[1.39,1.29]	[1.18,1.24]	[1.37,1.31]	[1.18,1.24]	3.26×10^{-4}	7.19×10^{-7}	43.64	171.99
Figure 12(d)	IV-(c)	—	2	6.1	[1.69,0.89]	[1.22,1.24]	[1.83,0.54]	[1.22,1.24]	7.09×10^{-2}	9.31×10^{-7}	42.24	249.01

---

[7] Stands for *second order statistic of the eigenvalues.*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Algorithm 4.1. Blind ToF imaging algorithm.

```
Input: Noisy ToF measurements $\mathbf{g}_{\mathbf{r}}$.
Kernel initialization: Compute $\boldsymbol{\varphi}_{\mathbf{r}}^{[0]}$ via (4.19).
for loop = 1 to max. iterations do
  for $j = 1$ to $j_{\max}$ do
    Construct the matrices in (4.14);
    Update $\mathbf{p}_{\mathbf{r}}^{[j]}, \mathbf{q}_{\mathbf{r}}^{[j]}$ by solving (4.17);
    if (4.20) holds then
      Terminate all loops;
    end if
  end for
  Calculate $\{\tau_{\mathbf{r}}[k], \Gamma_{\mathbf{r}}[k]\}_{k \in \mathbb{I}_K}$ using (4.12);
  Update $\boldsymbol{\varphi}_{\mathbf{r}}$ using (4.18);
  if (4.20) holds then
    Terminate all loops;
  end if
end for
Output: The SRF parameters $\{\tau_{\mathbf{r}}[k], \Gamma_{\mathbf{r}}[k]\}_{k \in \mathbb{I}_K}$ and the kernel $\boldsymbol{\varphi}_{\mathbf{r}}$.
```

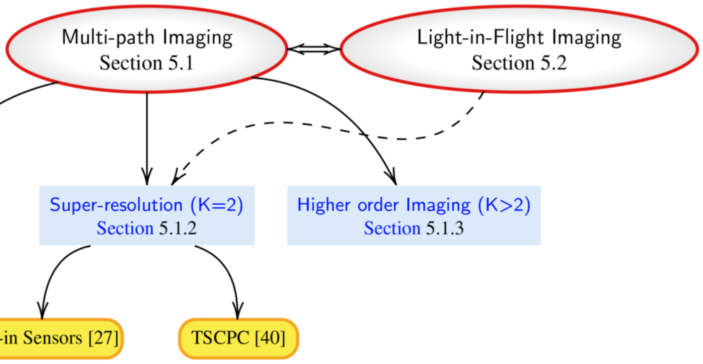


Figure 7. *Roadmap for hardware-based ToF imaging experiments. We demonstrate the performance of the proposed blind ToF imaging approach across various experimental setups and data modalities.*

5. Experiments.

Our experiments aim to validate the effectiveness of the blind ToF imaging approach, demonstrating that it performs comparably to methods with prior kernel calibration, which we use as a baseline for ground truth. Through a series of 10 experiments, we achieve robust 3D scene reconstruction across various experimental setups and ToF datasets without any kernel calibration. Figure 7 illustrates the different ToF imaging scenarios included in our experiments. We assess performance using the following quantitative metrics. MSE $\mathcal{E}(\boldsymbol{\Gamma}_{\mathbf{r}}, \widetilde{\boldsymbol{\Gamma}}_{\mathbf{r}})$ and $\mathcal{E}(\boldsymbol{\tau}_{\mathbf{r}}, \widetilde{\boldsymbol{\tau}}_{\mathbf{r}})$ are used to evaluate the 3D scene reconstruction accuracy. Peak signal-to-noise ratio (PSNR) is utilized to assess the accuracy of kernel estimation, i.e., $\mathsf{PSNR}(\boldsymbol{\varphi}_{\mathbf{r}}, \widetilde{\boldsymbol{\varphi}}_{\mathbf{r}}) = 10\log_{10}(\frac{\|\boldsymbol{\varphi}_{\mathbf{r}}\|_\infty^2}{\mathcal{E}(\boldsymbol{\varphi}_{\mathbf{r}}, \widetilde{\boldsymbol{\varphi}}_{\mathbf{r}})})$. We tabulate the experimental parameters and performance evaluation in Table 1.

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

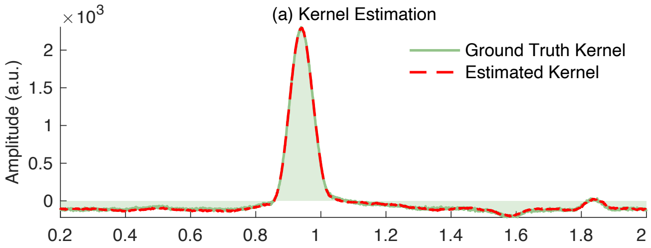
Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

5.1. Multipath imaging. Commercial ToF systems determine distance by amplitude modulation of a continuous wave. Although ToF sensors typically provide a single depth value per pixel, the actual scene might contain multiple depths, such as a semitransparent surface in front of a wall. This situation is known as a mixed pixel scenario [13, 35], where multiple light paths converge at the same pixel, resulting in a measured range that is a nonlinear mixture of the incoming paths.

To address the issue of multipath interference, one effective strategy involves emitting a custom code and capturing a series of demodulated values through successive electronic delays. Subsequently, a sparse deconvolution [13, 35] is applied to isolate a sequence of Diracs in the time profile, each corresponding to different light path depths and multipath combinations. We consider the results from this method as the ground truth for validating our blind recovery approach on these ToF datasets.

5.1.1. Diffuse imaging. We use the setup from [13, 35]. The scene consists of a placard reading "TIME OF FLIGHT" which is hidden by a diffusive semitranslucent sheet. Raw data comprising a $120 \times 120 \times 2976$ image tensor is acquired. Here, $N = 2976$ refers to the equidistant/uniform ToF measurements captured with sampling time $T = 70$ ps. The ToF measurements have been acquired with a custom ToF camera equipped with a PMD 19k-S3 sensor utilizing the architecture previously used in [5, 13, 35]. Both the illumination and the ToF pixels are modulated using the same binary M-sequences. This leads to a relatively narrow cross-correlation function in the time domain. The illumination control signal can be shifted with regard to the pixel control signal, thus shifting the resulting cross-correlation function. To contextualize our results, we first show the blind recovery of a single-pixel ToF data ($\mathbf{r} = [60\,60]^\top$) in Figure 8. The estimated kernel and reconstructed pixel measurements reach accuracy with $\mathsf{PSNR}(\boldsymbol{\varphi}_{\mathbf{r}}, \widetilde{\boldsymbol{\varphi}}_{\mathbf{r}}) = 43.24$ dB, $\mathcal{E}(\boldsymbol{\Gamma}_{\mathbf{r}}, \widetilde{\boldsymbol{\Gamma}}_{\mathbf{r}}) = 2.31 \times 10^{-5}$, and $\mathcal{E}(\boldsymbol{\tau}_{\mathbf{r}}, \widetilde{\boldsymbol{\tau}}_{\mathbf{r}}) = 1.87 \times 10^{-4}$. Then, we reconstruct the 3D scene from estimated time profiles by running the blind recovery


(a) Kernel Estimation

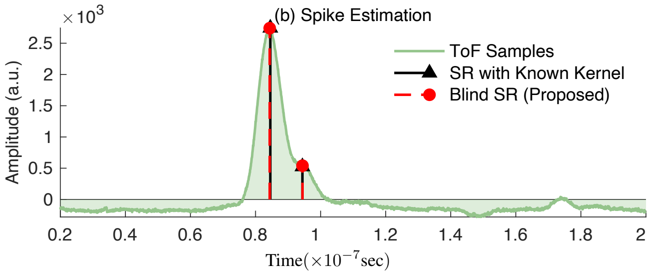

(b) Spike Estimation

Figure 8. *Blind recovery of single pixel ToF measurements* ($\mathbf{r} = [60\,60]^\top$, $K = 2$). (a) *Kernel reconstruction.* (b) *3D SRF estimation. The experimental setup is shown in Figure* 2.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

algorithm for each pixel measurements. This leads to indistinguishable results compared to the ground truth shown in Figure 1.

5.1.2. Super-resolution. In the context of multipath imaging, a natural and challenging scenario is to resolve closely located objects from time profiles. In order to validate the SRes capability of the proposed approach, we perform experiments on two different setups using lock-in sensors and TCSPC-TOF systems.

Lock-in sensors. The 3D scene comprises a mannequin head placed between a diffusive semitranslucent surface and a wall in the backdrop, as shown in Figure 2. By shifting the translucent surface, the interobject separation between the diffusive surface and the mannequin head is gradually reduced from 160, 130, 100 to 70 cm, respectively, with 30 cm reduction in each experiment. This equivalently leads to equidistant spike shifts as shown in Figure 9 (dashed lines). For this dataset, raw data comprising a $120 \times 120 \times 3968 \times 4$ image

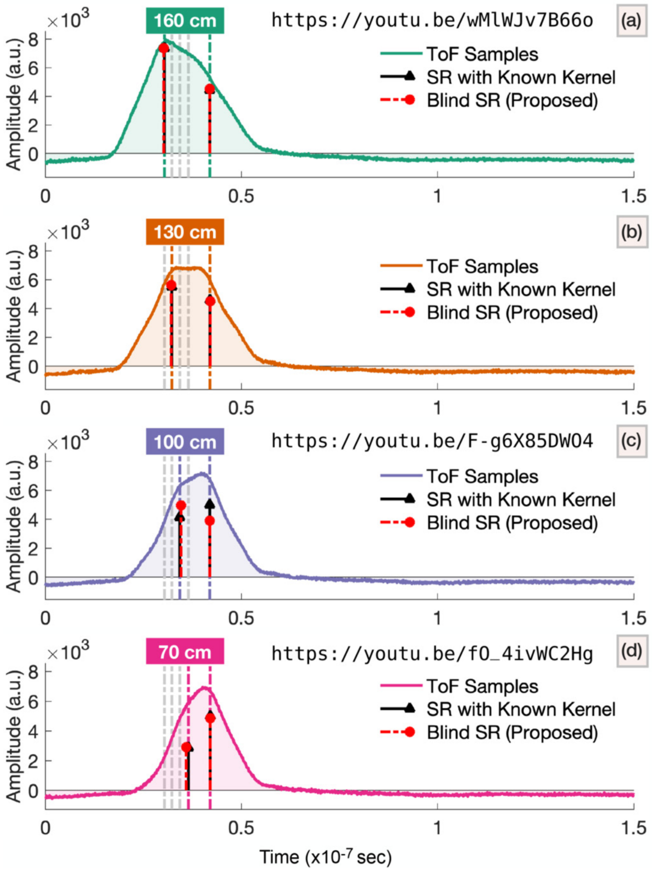


Figure 9. *Lock-in sensor based super-resolved ToF imaging. We benchmark the performance of the proposed method with kernel calibration. The scene consists of a mannequin head placed between a diffusive semitranslucent surface and a wall in the backdrop. By moving the diffusive surface, the interobject separation is uniformly reducing from* (a) 160, (b) 130, (c) 100, *to* (d) 70 *cm, respectively. The dashed lines represent the ground truth spikes. Our method super-resolves the interobject separation for* (a) 161, (b) 134, (c) 98, *and* (d) 72 *cm, which accurately matches the experimental setup in Figure* 2.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

tensor is acquired from a lock-in ToF sensor, for which $N = 3968$ refers to the number of ToF measurements recorded with sampling time $T = 96.15$ ps.

The blind SRF recoveries of a single-pixel ToF measurement ($\mathbf{r} = [60\,60]^\top, K = 2$) are presented in Figure 9. The translucent surface (front) is moving closer to the mannequin head (back), leading to the challenges of separating two objects. This can be intuitively observed from Figure 9(c) and (d), where one spike is invisible as the interobject distance is smaller than the kernel width. We validate the reconstructed 3D scene information by comparing the estimated interobject separation to the actual scene setups. The interobject separation is computed by $\Delta\widetilde{d}_{mk} = |\Delta\widetilde{\tau}_m - \Delta\widetilde{\tau}_k|c/2$, where $c = 3\times 10^8$m/s is the speed of light. Here, the Δd_{mk} refers to the results using the calibrated kernel. We outline the estimated results as follows: (1) 160 cm in Figure 9(a): $\Delta\widetilde{d}_{01} = 161$ and $\Delta d_{01} = 159$ cm; (2) 130 cm in Figure 9(b): $\Delta\widetilde{d}_{01} = 134$ and $\Delta d_{01} = 131$ cm; (3) 100 cm in Figure 9(c): $\Delta\widetilde{d}_{01} = 98$ and $\Delta d_{01} = 100$ cm; and (4) 70 cm in Figure 9(d): $\Delta\widetilde{d}_{01} = 72$ and $\Delta d_{01} = 69$ cm. This gives rise to the estimated target shift step with $27, 36, 26$ cm (blind recovery) and $28, 31, 31$ cm (known kernel). Compared to 30 cm shift difference from the actual scene setup, this demonstrates the SRes capability of the proposed blind recovery method. The 3D visualization of the object depth imaging is presented in Figure 10. We further reconstruct the 3D scene from the estimated time profiles by running the algorithm for all pixel measurements, as shown in Figure 11. Despite the challenging experimental setup, the proposed method super-resolves two objects up to a separation uncertainty of 5 cm, which is equivalent to resolving a separation of 0.33 ns in the time domain. This effectively demonstrates the SRes capability of the proposed method.

TCSPC. The experimental setup is borrowed from [14, 33] and comprises two retroreflecting corner cubes. In this setup, with one cube fixed, the interspacing with the second cube is

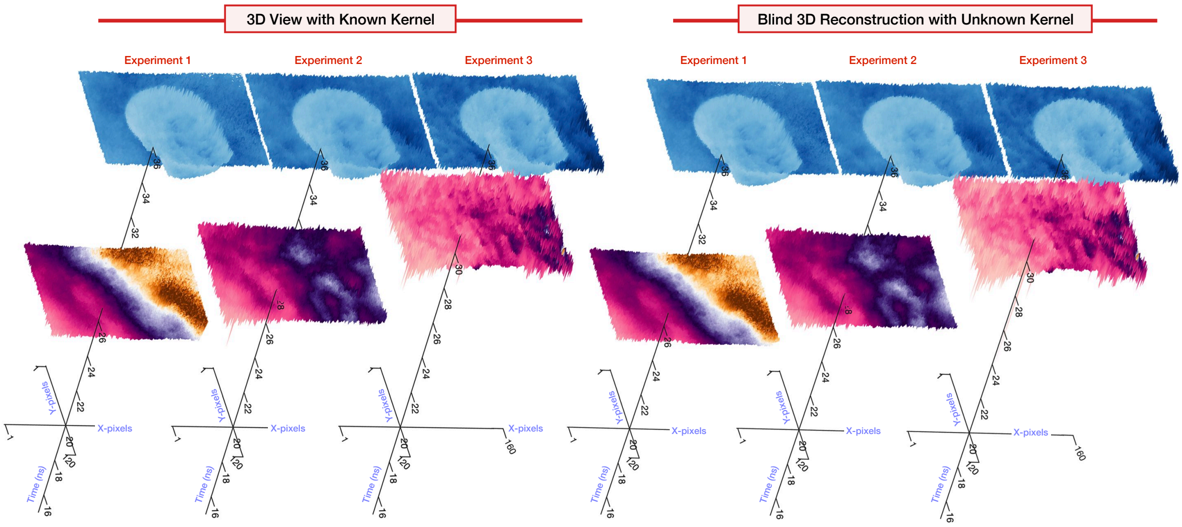


Figure 10. *Benchmarking blind ToF imaging with kernel calibration. The distance between the translucent surface and the mannequin head is reducing from* 160, 100, *to* 70 *cm, respectively, as shown in Figure* 2. *This requires algorithmic SRes capability.* (a) 3*D visualization of depth imaging with kernel calibration.* (b) 3*D visualization of blind depth imaging using the proposed blind recovery method.*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

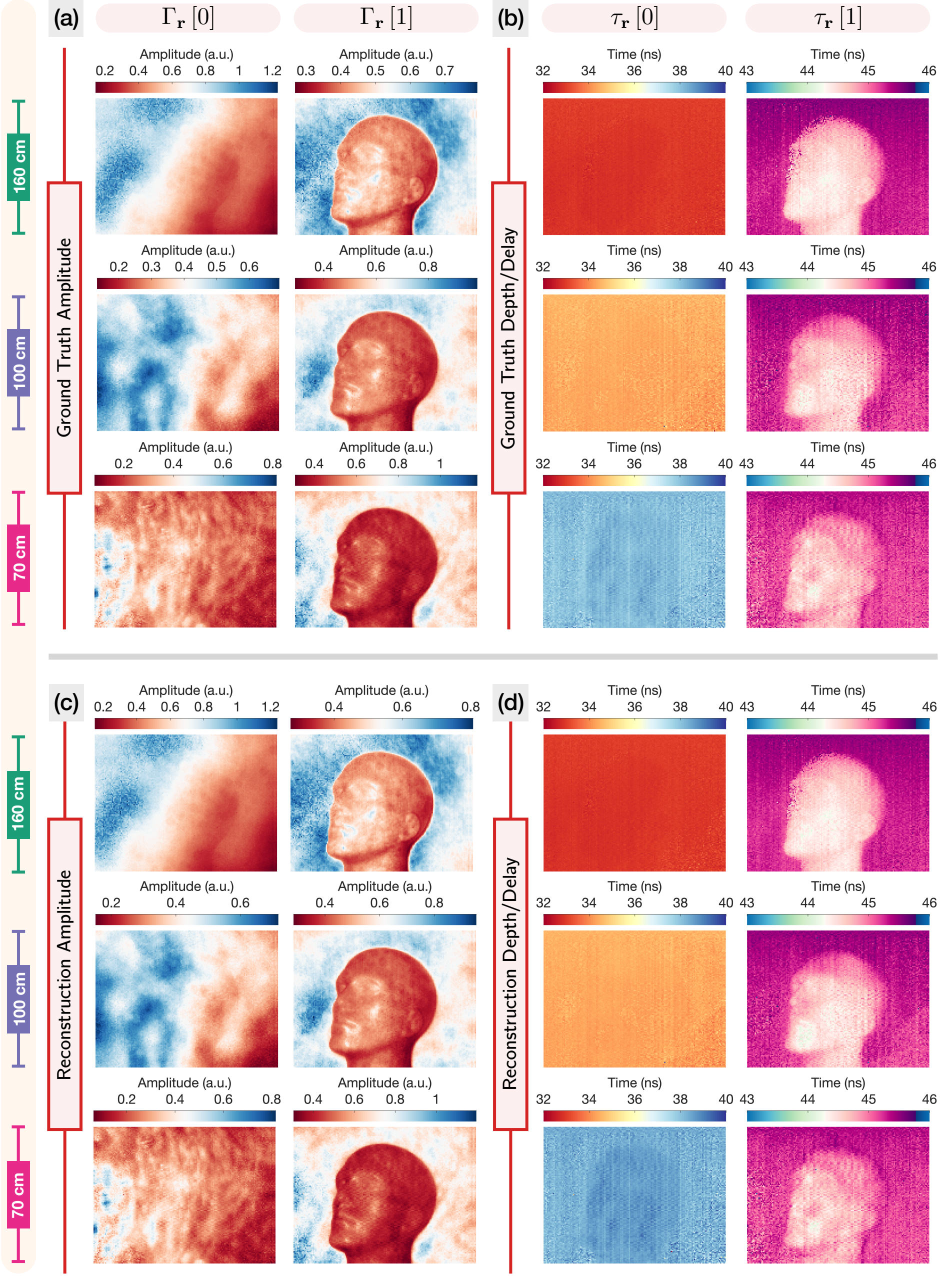


Figure 11. *Super-resolving object through a diffusive semitranslucent sheet.* (a) *and* (b) *are amplitude and depth imaging using known kernel;* (c) *and* (d) *are corresponding results utilizing our blind recovery method. The experimental setup is shown in Figure* 2*. Despite the challenges of small interobject separation, our method still achieves accurate scene reconstructions in all scenarios.*

reduced to create a scenario that requires algorithmic SRes [14]. For a corner cube, all beams, independent of incident angle, are reflected back in the original direction so the behavior is that of a perfect reflecting surface. In these experiments, the sampling step of the receiver was

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

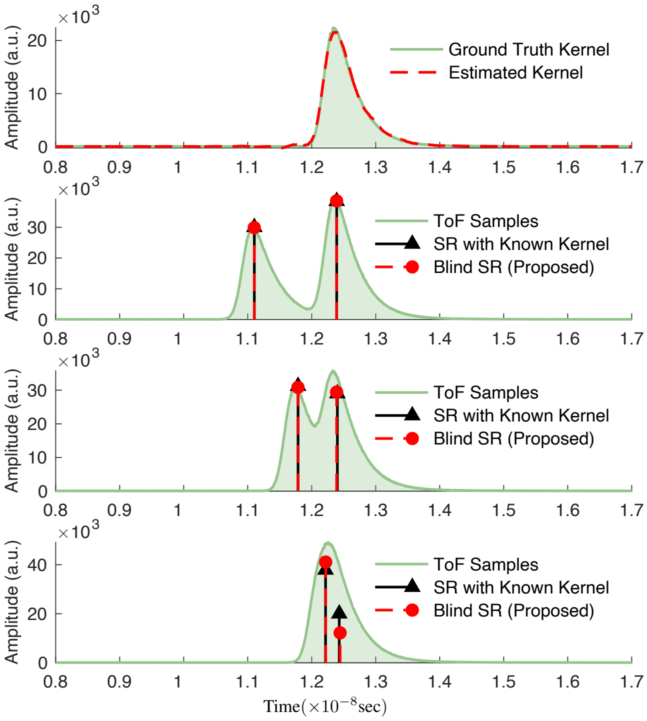


Figure 12. *LiDAR based super-resolved ToF imaging. The scene consists of two retro-reflecting cubes at a distance of* 330 *m from the TCSPC system. We show the blind recovery of the kernel and object depths at different interreflector separations. Our blind recovery method resolves two reflectors up to a resolution of* 13.47 *ps (equivalently* 0.20 *cm), showcasing the SRes capability of our method.*

$T = 6.1$ ps, and the collection time for each histogram was 30 seconds. We plot the estimated time-localized kernel and the raw measurements from three different interreflector separations in Figure 12(a)–(d), respectively. Similar to the scenarios in Figure 9, the task becomes more challenging as interreflector separation gets smaller. As shown in Figure 12, the reconstructed results are almost indistinguishable between the ground truth (obtained using [14]) and its recovery (see Table 1), which also matches the experimental setup and reported results [14]. It is noteworthy that the results in Figure 12 demonstrate the SRes of our method that resolves a separation of 13.47 ps in the time domain (equivalently 0.20 cm).

5.1.3. High-order imaging. This experiment is dedicated to pushing the limit of our method in the scenarios of high-order multipath imaging ($K = 3$). We use a calibrated scene with two translucent surfaces with a wall in the backdrop. The interobject separation is 1.8 and 2 meters, respectively, with a sampling step of $T = 70$ ps.

The blind recovery of two pixel measurements with coordinates $\mathbf{r} = [50\,70]^\top$ and $\mathbf{r} = [45\,60]^\top$ is plotted in Figure 13. It is worth noting that, compared to the data in Figures 8 and 9, the raw ToF measurements in Figure 13 are contaminated by higher levels of noise, creating algorithmic challenges for scene reconstruction. From the experimental results in Table 1, the estimated interobject separation from blind recovery is given by

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

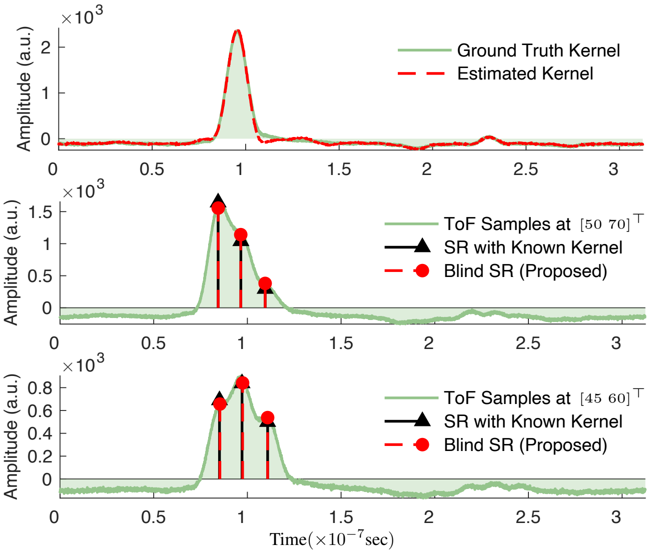


Figure 13. *Benchmarking blind ToF imaging for high-order reflections* ($K = 3$)*. Two translucent surfaces are placed in front of a wall with separation of* 1.8 *and* 2 *meters, respectively. The kernel and SRF recovery of two pixel measurements are plotted, showcasing the SRes capability of our method.*

1. pixel at $\mathbf{r} = [45\,60]^\top$, $\Delta\widetilde{d}_{01} = 1.81, \Delta\widetilde{d}_{12} = 1.95$ meters;
2. pixel at $\mathbf{r} = [50\,70]^\top$, $\Delta\widetilde{d}_{01} = 1.82, \Delta\widetilde{d}_{12} = 2.01$ meters.

This accurately matches the experimental setup (i.e., 1.8 and 2 meters) and the reported results in [8] that used the known kernel, corroborating the effectiveness of the proposed method.

5.2. Light-in-flight imaging.

In this experiment, the scene consists of a mannequin head located between a diffusive semitranslucent surface and a wall in the backdrop. The inter-object distance between the mannequin head and the wall is 160 cm, as demonstrated in Figures 2 and 10. For this dataset, raw data comprising a $120 \times 120 \times 3968$ image tensor is acquired from a lock-in ToF sensor with sampling time $T = 96.15$ ps. Using our custom ToF camera and blind recovery approach, we are able to visualize light sweep over the scene with multipath effects, as shown in Figure 14. In the early time-slices, the light first hits the diffusive sheet in Figure 14(k) (8.59 ns). The light then sweeps over the mannequin head on the scene from (o) (11.72 ns) to (t) (15.63 ns), and eventually hits the back wall in (y) (19.53 ns). The time slices correspond to the true geometry of the scene. The light sweep movies for inter-object separation of 160, 100 to 70 cm, respectively, can be visualized via the YouTube links provided in Table 2.

6. Conclusion.

This paper tackles a fundamental algorithmic challenge in optical time-of-flight (ToF) imaging. ToF imaging systems, which are active devices, illuminate scenes with light pulses or kernels and reconstruct them by capturing echoes of back-scattered light. This process enables the understanding of the 3D environment but introduces the challenge of sparse super-resolution (SRes), as the echoes exist at a finer time resolution than what conventional digital devices can measure. Additionally, the quality of SRes depends on the knowledge of the probing light pulses.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

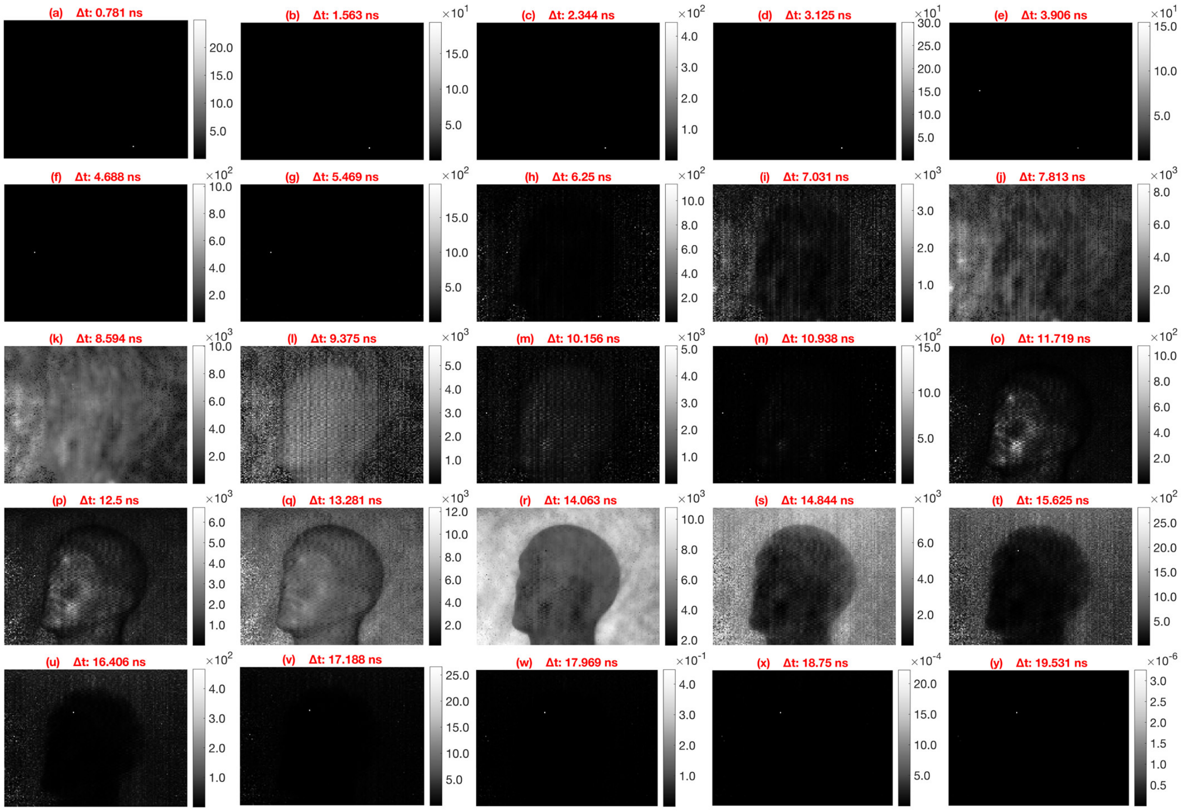


Figure 14. *LIF imaging: The scene consists of a mannequin head placed between a diffusive semitranslucent surface and a wall in the backdrop with a separation of* 160 *cm, as shown in Figure* 2*. Our method enables the visualization of the light sweep process without any kernel calibration. Light first impinges the diffusive surface in* (k)*, then sweeps over the mannequin head from* (o) *to* (t)*, and finally reaches the back wall as shown in* (y)*. We refer to Table* 2 *for visualization of light sweep movies corresponding to interobject separation of* 160, 100 *to* 70 *cm, respectively. These movies provide a visual interpretation of pulse propagation through the scene set up in Figure* 2*.*

Table 2
YouTube web links for light sweep movies.

Interobject separation	Web link (YouTube)
160 cm	https://youtu.be/wMlWJv7B66o
100 cm	https://youtu.be/F-g6X85DWO4
70 cm	https://youtu.be/fO_4ivWC2Hg

Departing from previous methods, this paper presents an algorithmic strategy that achieves sparse SRes without needing information about the emitted light pulses. Our algorithm effectively recovers sparse echoes modeled as continuous-time Dirac impulses, providing a more accurate representation of physical reality. Consequently, our approach eliminates the need for pulse calibration, offering a highly adaptable framework for blind ToF imaging that extends the limits of time resolution in 3D scene reconstruction. The validation of our approach

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

through extensive hardware experiments, encompassing a variety of scenarios and ToF modalities, demonstrates the empirical robustness and practical benefits of our method. Looking ahead, our algorithmic framework opens several avenues for future research, particularly in expanding this blind approach to additional contexts.

1. **Algorithmic frameworks.** Currently, we do not exploit the fact that all pixels utilize the same pulse. This scenario has been explored under the theme of multichannel sparse blind deconvolution [36, 58] and applying similar priors to our case is highly pertinent. Moreover, considering the blind recovery scenario where each spike is filtered through a distorted version of the original pulse [6] can pave the way for new application areas.
2. **New sensing pipelines.** Exploring nonlinear sensing modalities leads to the emergence of new classes of inverse problems and provides clear practical advantages over traditional pointwise sampling. In this context, we have identified two potential areas for development: (i) one-bit sensing [8], which offers a low-complexity implementation, and (ii) the unlimited sensing framework [4, 11, 12], which simultaneously provides high dynamic range and high digital resolution.
3. **Application areas.** Our methodology is applicable to a range of challenges including fluorescence lifetime imaging [5], looking-around-the-corners [55], and imaging through scattering [61], all of which are significant for practical implementations. Moreover, beyond the realm of optical ToF methods, our approach has potential advantages for related ToF systems such as seismic imaging [36] and radar technologies [22].

REFERENCES

[1] N. Abramson, *Light-in-flight recording by holography*, Opt. Lett., 3 (1978), pp. 121–123, https://doi.org/10.1364/OL.3.000121.

[2] A. Adam, C. Dann, O. Yair, S. Mazor, and S. Nowozin, *Bayesian time-of-flight for realtime shape, illumination and albedo*, IEEE Trans. Pattern Anal. Mach. Intell., 39 (2017), pp. 851–864, https://doi.org/10.1109/TPAMI.2016.2567379.

[3] A. J. Bell and T. J. Sejnowski, *An information-maximization approach to blind separation and blind deconvolution*, Neural Comput., 7 (1995), pp. 1129–1159, https://doi.org/10.1162/neco.1995.7.6.1129.

[4] A. Bhandari, *Back in the US-SR: Unlimited sampling and sparse super-resolution with its hardware validation*, IEEE Signal Process. Lett., 29 (2022), pp. 1047–1051, https://doi.org/10.1109/LSP.2022.3161865.

[5] A. Bhandari, C. Barsi, and R. Raskar, *Blind and reference-free fluorescence lifetime estimation via consumer time-of-flight sensors*, Optica, 2 (2015), 965, https://doi.org/10.1364/OPTICA.2.000965.

[6] A. Bhandari and T. Blu, *FRI sampling and time-varying pulses: Some theory and four short stories*, in Proceedings of the International Conference on Acoustics, Speech and Signal Processing, IEEE, 2017.

[7] A. Bhandari, A. Bourquard, and R. Raskar, *Sampling without time: Recovering echoes of light via temporal phase retrieval*, in Proceedings of the International Conference on Acoustics, Speech and Signal Processing, IEEE, 2017.

[8] A. Bhandari, M. H. Conde, and O. Loffeld, *One-bit time-resolved imaging*, IEEE Trans. Pattern Anal. Mach. Intell., 42 (2020), pp. 1630–1641, https://doi.org/10.1109/TPAMI.2020.2986950.

[9] A. Bhandari, A. Kadambi, and R. Raskar, *Sparse linear operator identification without sparse regularization? Applications to mixed pixel problem in time-of-flight range imaging*, in Proceedings of the International Conference on Acoustics, Speech and Signal Processing, IEEE, 2014, https://doi.org/10.1109/ICASSP.2014.6853619.

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

[10] A. Bhandari, A. Kadambi, and R. Raskar, *Computational Imaging*, MIT Press, Cambridge, MA, 2022, https://imagingtext.github.io/.

[11] A. Bhandari, F. Krahmer, and T. Poskitt, *Unlimited sampling from theory to practice: Fourier-Prony recovery and prototype ADC*, IEEE Trans. Signal Process., 70 (2021), pp. 1131–1141, https://doi.org/10.1109/TSP.2021.3113497.

[12] A. Bhandari, F. Krahmer, and R. Raskar, *On unlimited sampling and reconstruction*, IEEE Trans. Signal Process., 69 (2020), pp. 3827–3839, https://doi.org/10.1109/TSP.2020.3041955.

[13] A. Bhandari and R. Raskar, *Signal processing for time-of-flight imaging sensors: An introduction to inverse problems in computational 3-d imaging*, IEEE Signal Process. Mag., 33 (2016), pp. 45–58, https://doi.org/10.1109/MSP.2016.2582218.

[14] A. Bhandari, A. M. Wallace, and R. Raskar, *Super-resolved time-of-flight sensing via FRI sampling theory*, in Proceedings of the International Conference on Acoustics, Speech and Signal Processing, IEEE, 2016.

[15] T. Blu, H. Bay, and M. Unser, *A new high-resolution processing method for the deconvolution of optical coherence tomography signals*, in Proceedings of the International Symposium on Biomedical Imaging, IEEE, 2002.

[16] C. A. Bouman, *Foundations of Computational Imaging: A Model-Based Approach*, SIAM, Philadelphia, 2022, https://doi.org/10.1137/1.9781611977134.

[17] E. J. Candès and C. Fernandez-Granda, *Towards a mathematical theory of super-resolution*, Comm. Pure Appl. Math., 67 (2013), pp. 906–956, https://doi.org/10.1002/cpa.21455.

[18] J. Castorena and C. D. Creusere, *Sampling of time-resolved full-waveform LIDAR signals at sub-Nyquist rates*, IEEE Trans. Geosci. Remote Sens., 53 (2015), pp. 3791–3802, https://doi.org/10.1109/TGRS.2014.2383839.

[19] P. Catala, V. Duval, and G. Peyré, *A low-rank approach to off-the-grid sparse superresolution*, SIAM J. Imaging Sci., 12 (2019), pp. 1464–1500, https://doi.org/10.1137/19M124071X.

[20] J. F. Claerbout and F. Muir, *Robust modeling with erratic data*, Geophysics, 38 (1973), pp. 826–844, https://doi.org/10.1190/1.1440378.

[21] M. H. Conde, *A material-sensing time-of-flight camera*, IEEE Sensors Lett., 4 (2020), pp. 1–4, https://doi.org/10.1109/LSENS.2020.3005042.

[22] R. J. P. De Figueiredo and C.-L. Hu, *Waveform feature extraction based on Tauberian approximation*, IEEE Trans. Pattern Anal. Mach. Intell., 4 (1982), pp. 105–116, https://doi.org/10.1109/TPAMI.1982.4767214.

[23] Q. Denoyelle, V. Duval, G. Peyré, and E. Soubies, *The sliding Frank-Wolfe algorithm and its application to super-resolution microscopy*, Inverse Problems, 36 (2019), 014001, https://doi.org/10.1088/1361-6420/ab2a29.

[24] D. L. Donoho, *Superresolution via sparsity constraints*, SIAM J. Math. Anal., 23 (1992), pp. 1309–1331, https://doi.org/10.1137/0523074.

[25] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, *Single-pixel imaging via compressive sampling*, IEEE Signal Process. Mag., 25 (2008), pp. 83–91, https://doi.org/10.1109/MSP.2007.914730.

[26] M. Elad, *Sparse and Redundant Representations*, Springer, New York, 2010.

[27] S. Foix, G. Alenya, and C. Torras, *Lock-in time-of-flight (ToF) cameras: A survey*, IEEE Sensors J., 11 (2011), pp. 1917–1926, https://doi.org/10.1109/JSEN.2010.2101060.

[28] L. Gao, J. Liang, C. Li, and L. V. Wang, *Single-shot compressed ultrafast photography at one hundred billion frames per second*, Nature, 516 (2014), pp. 74–77, https://doi.org/10.1038/nature14005.

[29] 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-fight imaging*, Nat. Commun., 6 (2015), https://doi.org/10.1038/ncomms7021.

[30] R. Guo, Y. Li, T. Blu, and H. Zhao, *Vector-FRI recovery of multi-sensor measurements*, IEEE Trans. Signal Process., 70 (2022), pp. 4369–4380, https://doi.org/10.1109/TSP.2022.3204402.

[31] Z. He, A. Cichocki, S. Xie, and K. Choi, *Detecting the number of clusters in n-way probabilistic clustering*, IEEE Trans. Pattern Anal. Mach. Intell., 32 (2010), pp. 2006–2021, https://doi.org/10.1109/TPAMI.2010.15.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

[32] F. Heide, M. B. Hullin, J. Gregson, and W. Heidrich, *Low-budget transient imaging using photonic mixer devices*, ACM Trans. Graphics, 32 (2013), pp. 1–10, https://doi.org/10.1145/2461912.2461945.

[33] S. Hernandez-Marin, A. Wallace, and G. Gibson, *Bayesian analysis of lidar signals with multiple returns*, IEEE Trans. Pattern Anal. Mach. Intell., 29 (2007), pp. 2170–2180, https://doi.org/10.1109/TPAMI.2007.1122.

[34] A. Jarabo, B. Masia, J. Marco, and D. Gutierrez, *Recent Advances in Transient Imaging: A Computer Graphics and Vision Perspective*, Vol. 1, Elsevier, Amsterdam, 2017.

[35] 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 Trans. Graphics, 32 (2013), pp. 1–10, https://doi.org/10.1145/2508363.2508428.

[36] N. Kazemi and M. D. Sacchi, *Sparse multichannel blind deconvolution*, GeoPhys., 79 (2014), pp. V143–V152, https://doi.org/10.1190/geo2013-0465.1.

[37] A. Kirmani, D. Venkatraman, D. Shin, A. Colaco, F. N. C. Wong, J. H. Shapiro, and V. K. Goyal, *First-photon imaging*, Science, 343 (2014), pp. 58–61, https://doi.org/10.1126/science.1246775.

[38] H.-W. Kuo, Y. Zhang, Y. Lau, and J. Wright, *Geometry and symmetry in short-and-sparse deconvolution*, SIAM J. Math. Data Sci., 2 (2020), pp. 216–245, https://doi.org/10.1137/19M1237569.

[39] X. Li, S. Ling, T. Strohmer, and K. Wei, *Rapid, robust, and reliable blind deconvolution via nonconvex optimization*, Appl. Comput. Harmon. Anal., 47 (2019), pp. 893–934, https://doi.org/10.1016/j.acha.2018.01.001.

[40] S. Pellegrini, G. S. Buller, J. M. Smith, A. M. Wallace, and S. Cova, *Laser-based distance measurement using picosecond resolution time-correlated single-photon counting*, Meas. Sci. Technol., 11 (2000), pp. 712–716, https://doi.org/10.1088/0957-0233/11/6/314.

[41] D. Perrone and P. Favaro, *A clearer picture of total variation blind deconvolution*, IEEE Trans. Pattern Anal. Mach. Intell., 38 (2016), pp. 1041–1055, https://doi.org/10.1109/TPAMI.2015.2477819.

[42] C. Poon and G. Peyré, *Multidimensional sparse super-resolution*, SIAM J. Math. Anal., 51 (2019), pp. 1–44, https://doi.org/10.1137/17M1147822.

[43] H. Qiao, J. Lin, Y. Liu, M. B. Hullin, and Q. Dai, *Resolving transient time profile in ToF imaging via log-sum sparse regularization*, Opt. Lett., 40 (2015), pp. 918, https://doi.org/10.1364/OL.40.000918.

[44] A. Redo-Sanchez, B. Heshmat, A. Aghasi, S. Naqvi, M. Zhang, J. Romberg, and R. Raskar, *Terahertz time-gated spectral imaging for content extraction through layered structures*, Nat. Commun., 7 (2016), https://doi.org/10.1038/ncomms12665.

[45] A. Repetti, M. Q. Pham, L. Duval, E. Chouzenoux, and J.-C. Pesquet, *Euclid in a taxicab: Sparse blind deconvolution with smoothed, ℓ_1/ℓ_2 regularization*, IEEE Signal Process. Lett., 22 (2015), pp. 539–543.

[46] G. Satat, B. Heshmat, C. Barsi, D. Raviv, O. Chen, M. G. Bawendi, and R. Raskar, *Locating and classifying fluorescent tags behind turbid layers using time-resolved inversion*, Nat. Commun., 6 (2015), https://doi.org/10.1038/ncomms7796.

[47] C. S. Seelamantula and S. Mulleti, *Super-resolution reconstruction in frequency-domain optical-coherence tomography using the finite-rate-of-innovation principle*, IEEE Trans. Signal Process., 62 (2014), pp. 5020–5029, https://doi.org/10.1109/TSP.2014.2340811.

[48] P. Sen, B. Chen, G. Garg, S. R. Marschner, M. Horowitz, M. Levoy, and H. P. A. Lensch, *Dual photography*, in Proceedings of SIGGRAPH, ACM, 2005.

[49] D. Shin, F. Xu, F. N. C. Wong, J. H. Shapiro, and V. K. Goyal, *Computational multi-depth single-photon imaging*, Opt. Express, 24 (2016), 1873, https://doi.org/10.1364/OE.24.001873.

[50] G. Strang and G. Fix, *A Fourier Analysis of the Finite Element Variational Method*, Springer, Berlin, 2011.

[51] *First M*87 *Event Horizon Telescope results.* IV. *Imaging the central supermassive black hole*, Astrophys. J. Lett., 875 (2019).

[52] M. O. Toole, F. Heide, D. B. Lindell, K. Zang, S. Diamond, and G. Wetzstein, *Reconstructing transient images from single-photon sensors*, in Proceedings of the International Conference on Computer Vision and Pattern Recognition, IEEE, 2017.

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 06/20/25 to 146.179.86.172 . Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/terms-privacy

[53] R. Tur, Y. C. Eldar, and Z. Friedman, *Innovation rate sampling of pulse streams with application to ultrasound imaging*, IEEE Trans. Signal Process., 59 (2011), pp. 1827–1842, https://doi.org/10.1109/TSP.2011.2105480.

[54] J. A. Urigen, T. Blu, and P. L. Dragotti, *FRI sampling with arbitrary kernels*, IEEE Trans. Signal Process., 61 (2013), pp. 5310–5323, https://doi.org/10.1109/TSP.2013.2278152.

[55] A. Velten, T. Willwacher, O. Gupta, A. Veeraraghavan, M. G. Bawendi, and R. Raskar, *Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging*, Nat. Commun., 3 (2012), https://doi.org/10.1038/ncomms1747.

[56] A. Velten, D. Wu, A. Jarabo, B. Masia, C. Barsi, C. Joshi, E. Lawson, M. Bawendi, D. Gutierrez, and R. Raskar, *Femto-photography: Capturing and visualizing the propagation of light*, ACM Trans. Graphics, 32 (2013), pp. 1–8, https://doi.org/10.1145/2461912.2461928.

[57] A. Velten, D. Wu, B. Masia, A. Jarabo, C. Barsi, C. Joshi, E. Lawson, M. Bawendi, D. Gutierrez, and R. Raskar, *Imaging the propagation of light through scenes at picosecond resolution*, Commun. ACM, 59 (2016), pp. 79–86, https://doi.org/10.1145/2975165.

[58] L. Wang, Q. Zhao, J. Gao, Z. Xu, M. Fehler, and X. Jiang, *Seismic sparse-spike deconvolution via Toeplitz-sparse matrix factorization*, GeoPhys., 81 (2016), pp. V169–V182, https://doi.org/10.1190/geo2015-0151.1.

[59] W. Wang, J. Li, and H. Ji, *l_1-norm regularization for short-and-sparse blind deconvolution: Point source separability and region selection*, SIAM J. Imaging Sci., 15 (2022), pp. 1345–1372.

[60] C. Wu, J. Liu, X. Huang, Z.-P. Li, C. Yu, J.-T. Ye, J. Zhang, Q. Zhang, X. Dou, V. K. Goyal, F. Xu, and J.-W. Pan, *Non-line-of-sight imaging over* 1.43 *km*, Proc. Nat. Acad. Sci. USA, 118 (2021).

[61] D. Wu, M. O'Toole, A. Velten, A. Agrawal, and R. Raskar, *Decomposing global light transport using time of flight imaging*, in Proceedings of the International Conference on Computer Vision and Pattern Recognition, IEEE, 2012.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.