Nonuniform Sampling of Echoes of Light

Miguel Heredia Conde[†], Ayush Bhandari[‡] and Otmar Loffeld[†]

[†]Center for Sensorsystems (ZESS), University of Siegen, Paul-Bonatz-Straße 9-11, 57076 Siegen, Germany.
[‡]Department of Electrical and Electronic Engineering, Imperial College London, London SW7 2AZ, U.K.

Emails: heredia@zess.uni-siegen.de • a.bhandari@imperial.ac.uk • loffeld@zess.uni-siegen.de

Abstract—Recent work on time-resolved imaging (TRI) has shown that real-world scenes can often be explained by means of sparse time-domain responses. Direct sampling of such scene responses in the time-domain requires exorbitant sampling rates posing a practical bottleneck. An alternative approach uses sinusoidal illumination whereby the phase difference encodes time delays. When working with multiple frequencies this allows for directly sampling the Fourier spectrum of sparse scene responses and this is made possible using a Time-of-Flight (ToF) camera. Due to hardware restrictions, creating an equidistant set of frequencies is challenging. In this paper we adopt a nonuniform sampling architecture and propose an extension of the real-valued IAA (RIAA) algorithm which sequentially estimates the sparse components. Experimental validation with both synthetic and real data acquired with a ToF sensor confirms the feasibility of the proposed approach. This leverages the requirements on the hardware and paves the way for accurate scene response sensing with low-cost ToF sensors.

I. INTRODUCTION

Time-resolved imaging (TRI) is an emerging area of research which exploits the knowledge of time-delays resulting from the interaction between light and the scene. This form of imaging offers a new way to see the world. Some examples include non-line-of-sight imaging [1], [2], multi-depth imaging [3]–[9] and low-cost bio-imaging [10]. Research on TRI has unveiled that real-world scenes lead to continuous-time sparse profiles. For any pixel of a TRI system, the scene response it observes may be modeled as

$$h(t) = \sum_{k=1}^{K} \Gamma_k \delta(t - t_k), \qquad t_k = 2d_k/c, \qquad (1)$$

where t_k denotes the *time-of-flight* (ToF) or the echo generated by an object at distance d_k from the sensor and Γ_k is the reflectivity. In real-world scenes K is finite and typically low[1]. Clearly, sampling such scenes requires exorbitant samples rates, mainly because (1) is a non-bandlimited function. Alternatively, sampling the Fourier transform of $h(t)$ leads to a parametric representation—a weighted sum of cisoids (cf. measurements in Fig. 1). Estimating parameters of sum of cisoids is a well studied in the field of spectral estimation theory. Technologically, this sampling in Fourier domain is accomplished using time-of-flight (ToF) sensors [11].

Since the advent of low-cost ToF-based depth cameras, mass-produced and oriented to the entertainment market, the field of application of this technology has undergone an impressive expansion. These and similar sensors have triggered a considerable amount of novel research on 3D sensing, indoor mapping, robot navigation, SLAM, etc. Despite the large improvements in terms of accuracy and resolution, ToF sensors suffer from a number of known disadvantages that preclude their application in real-world environments.

ToF cameras probe the scene with modulated light and compute a single depth value per pixel from the received scene reflection. It is clear that the underlying hypothesis is that each pixel receives a single bounce, arising from a single scatterer in the scene. If this does not hold and the light signal reaching the ToF pixel is the sum of two or more reflected signals following different return paths the problem of
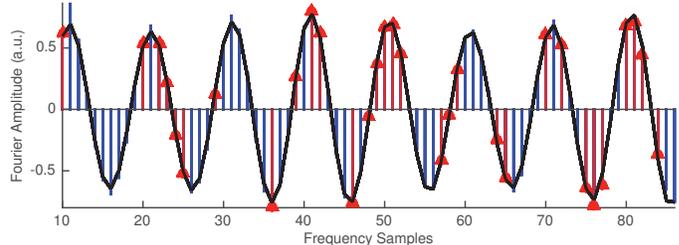
Fig. 1: Original uniformly-sampled raw data from a ToF pixel in frequency domain (blue stems) and data reconstruction using the parameters estimated by the nested-RIAA approach proposed in this paper (black curve) from a nonuniform sub-set of data. Only 40% samples were considered (red stems). The mean squared error in reconstruction is 0.0017.

multi-path interference or MPI arises and the depth estimate given by the ToF sensor will be erroneous. In order to maximize signal-to-noise ratio, most ToF systems [12] operate with sinusoidal illumination[2].

Under the single-bounce hypothesis, that is, $h(t) = \Gamma_0 \delta(t - t_0)$, a single frequency suffices to estimate 2 unknowns $\{\Gamma_0, t_0\}$. This is efficiently implemented on the hardware using the **Four Bucket Method** reviewed in Section III. Clearly, when $K > 1$, multiple echoes become indistinguishable with a single frequency setup. To overcome this problem, multi-frequency approaches have been developed in literature [5], [6]. These approaches rely on a dense uniform sampling in frequency domain and require customized hardware.

Here, we advocate a different idea— use of nonuniform frequency spacing which leads to nonuniform sampling in the Fourier domain. There are several reasons that favor this approach.

- In practice, periodic signals are constructed by frequency division of a single clock frequency. The implication is that some frequencies can be achieved exactly, while some others cannot.
- Depending on the quality of the electronics, additional jitter may appear due to electromagnetic interference and crosstalk leading to imperfections in uniform sampling lattices.
- From a hardware perspective, lower frequencies tend to be more stable than higher frequencies.

Motivated by the practicalities around the ToF hardware, in this paper we propose a new approach that is based on nonuniform sampling in the Fourier domain. The key idea is to resolve multiple echoes using a set of nonuniformly distributed frequencies. By means of simulations using both synthetic and real data from a ToF camera we show that multiple echoes can be estimated from very reduced sets of nonuniform samples in frequency domain. Fig. 1 illustrates the performance of the proposed approach, which is able to estimate

---

[1]This is because inter-reflections follow the inverse-square law.

[2]In ToF jargon, this is known as Amplitude Modulated Continuous Wave (AMCW) mode. Furthermore, the modulation is assumed to be quasi-sinusoidal, being the effects of eventual harmonic distortion often calibrated *a posteriori* in depth domain.

the unknown parameters in (1) from few nonuniform samples (in red), thus enabling an accurate parametric representation of the data (in black).

II. RELATED WORK

In this work we are interested in estimating several paths per pixel, thus eliminating the need for a subsequent computationally-expensive multipath-removal procedure. Furthermore, we restrict our attention exclusively to *reflective* multipath, arising from translucent objects and shiny surfaces, such as floors, mirrors, whiteboards, walls, windows, tables, metallic surfaces, etc. For $K = 2$, [4], [8] provide closed-form solutions. Alternatively, in [7] the multiple paths are retrieved by means of an optimization process. The methods in [7], [8], and [4] cope with only two paths per pixel and require two, three and five frequencies, respectively. If more than two bounces interfere within the same pixel, none of the previous approaches apply. Note that if the phase or depth domain is finely discretized, the few targets producing the MPI can be modeled as a *sparse* vector of reflectances and the multipath estimation problem can be attacked from a *compressive sensing* (CS) perspective [13]–[15]. This is the focus adopted in [5], where this sparse vector is recovered from partial Fourier measurements in a classical CS framework. Unfortunately, the method seems to require a large number of modulation frequencies in practice (77 for *known* sparsity of 3). A more feasible approach is provided in [6], where k interfering paths are estimated from $2k + 1$ frequency measurements in a closed-form manner. In an evaluation with real data from an Xbox One sensor this method required 21 measurements to separate two paths. The frequency-domain framework in [3] allows separating multiple paths at the cost of an unaffordable modulation bandwidth of the illumination system (e. g., $10\,\text{GHz}$ for $3.6\,\text{cm}$ depth resolution).

III. FROM ToF SENSORS TO FOURIER SAMPLES

We have already stated that the scene response function in (1) is to be sensed in Fourier domain by means of a ToF camera. Now it remains to clarify how a ToF camera gathers these Fourier measurements [5]. ToF cameras are active devices, that is, they emit light onto the scene. The emitted light is modulated in amplitude and, in order to perform Fourier sensing, this modulation will be sinusoidal. Thus the signals the camera emits $s(t)$ and receives $r(t)$ are (under the single-path assumption) of the shape

$$s(t) = 1 + s_0 \cos(\omega t)$$
$$r(t) = \Gamma\left(1 + s_0 \cos(\omega t - \phi)\right), \qquad \phi = \omega t_0 \tag{2}$$

where s_0 is the modulation depth and Γ the amplitude of the reflected signal, which is delayed by a phase ϕ w.r.t. $s(t)$. The distance between camera and reflector is then $d = c\phi/(2\omega)$, where c is the speed of light. Furthermore, the integration process at the ToF pixels is also regulated, typically by the same signal used to modulate the light. Consequently each pixel behaves as a homodyne detector and its measurements can be explained as samples of the cross-correlation between the reference signal and the received signal:

$$c_\omega[q] = (s \star r)(\tau_q) = \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} s^*(t + \tau_q)r(t)dt$$
$$= \Gamma\left(1 + \frac{s_0^2}{2}\cos(\omega\tau_q + \phi)\right). \tag{3}$$

Using the well-known *Four Bucket Method* both unknown parameters, Γ and ϕ, can be estimated from four samples of $c_\omega[q]$ at $\tau_q = \pi q/(2\omega)$, $q = 0, \dots, 3$. This is done by constructing a complex

number, $z_\omega = (c_\omega[0] - c_\omega[2]) + j(c_\omega[3] - c_\omega[1])$. The unknowns Γ and $\phi = \omega t_0$ are estimated using,

$$\widetilde{\Gamma} = |z_\omega|/s_0^2 \qquad \text{and} \qquad \widetilde{\phi}_0 = \angle z_\omega. \tag{4}$$

Thus the associated *complex measurement* reads:

$$y(\omega) = \widetilde{\Gamma}e^{j\widetilde{\phi}\omega}, \tag{5}$$

which is indeed the Fourier representation of a spike.

IV. TIME-OF-FLIGHT SENSING AND SPECTRAL ESTIMATION

As sketched in Section I, our goal is to estimate more than a single depth per pixel of a ToF camera operating under MPI conditions. To this end, measurements are carried out probing the scene with light signals in AMCW mode. Each probing signal is considered to be quasi-sinusoidal, with negligible harmonic distortion. Differently from prior work, where the measurements were obtained according to a uniform grid in frequency domain, we shift to a more general setting, where the sampling points in frequency domain are arbitrarily, *e.g.*, pseudo-randomly distributed. This is a physically-motivated scenario and enables a better adaptation to the specific ToF hardware capabilities by eliminating the artificial uniform-sampling restriction.

For completeness, note that multipath can be classified between *diffuse* and *reflective*. The first occurs due to light scattering, *e.g.*, due to Lambertian-reflective objects that are too close to the camera, in combination with low-quality optics. Scattering media, such as turbid water or translucent objects, may also produce diffuse multipath. In this work we restrict our attention to the second type, which arises from strong secondary reflections, *e.g.*, caused by shiny floors or walls, which produce a secondary illumination front for the rest of the scene. In this scenario, and supposing that we deal with K interfering paths, being K a low number, typically $K \leq 3$, the aim is to estimate a set of parameters $\{\Gamma_k, d_k\}_{k=1}^{K}$, which uniquely define the scene response function [11], [16], which in time domain can be written as in (1).

As pointed out in [11], [16] and references therein, the measurements of a ToF pixel operating in AMCW mode with sinusoidal modulation can be seen as samples of the Fourier transform of $h(t)$, formally:

$$y(t,\omega) = \frac{1}{2}e^{j\omega t}\sum_{k=1}^{K}\Gamma_k e^{+j\omega t_k} = \frac{1}{2}e^{j\omega t}\hat{h}^*(\omega),$$
$$\text{with } \hat{h}(\omega) = \sum_{k=1}^{K}\Gamma_k e^{-j\omega t_k} \tag{6}$$

where $\hat{h}(\omega)$ denotes the Fourier transform of $h(t)$. Both parameters t and ω can be adjusted in a ToF camera in order to obtain the desired measurements $y(t,\omega)$, the latter by setting the modulation frequency and the former by adjusting the delay between the signals for light modulation and pixel demodulation. In mono-frequency ToF ω is fixed to a frequency that is upper constrained by either the minimum unambiguous range or by bandwidth limitations of the illumination system and the ToF pixels and measurements are performed in time delay, i.e., phase shift, domain. Differently, in the case of multi-frequency ToF measurements are gathered in frequency domain for different values of ω, while t is left constant, eventually zero after appropriate calibration. In the latter case the measurements are directly the (conjugated) Fourier coefficients $y(\omega) = \hat{h}^*(\omega)$.

Let $\{\omega_i\}_{i=1}^{m}$ be the set of frequencies at which measurements are acquired and $\{y_i\}_{i=1}^{m}$ the set of measurements $y_i = y(\omega_i)$. The task of resolving the unknown parameters $\{\Gamma_k, d_k\}_{k=1}^{K}$ from the set of spectral samples $\{y_i\}_{i=1}^{m}$ is a classical spectral estimation problem.

To date the literature on ToF MPI estimation from frequency measurements used a regular sampling grid, that is a set of *Fourier*
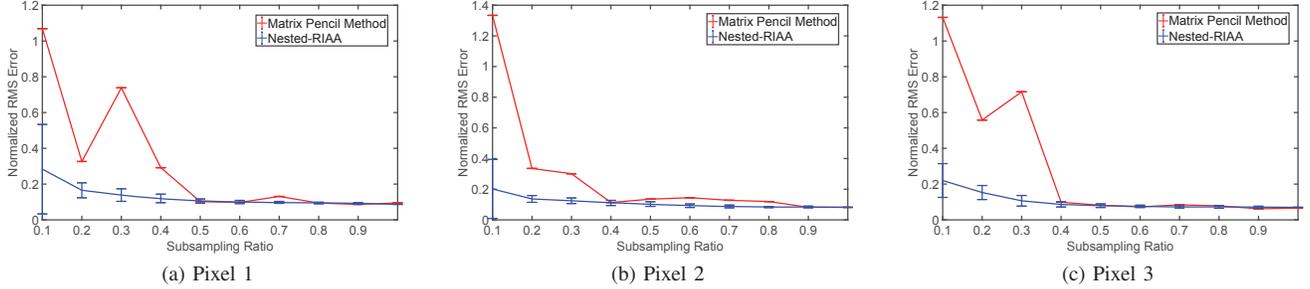
Fig. 2: Normalized RMSE of the reconstructed data using the matrix pencil method and uniform sampling in frequency domain (red line) and using the nested RIAA with nonuniform sampling in frequency domain. Mean values and standard deviations were obtained over 64 experimental runs with randomly-generated nonuniform sampling patterns. The data was collected from three different ToF pixels.
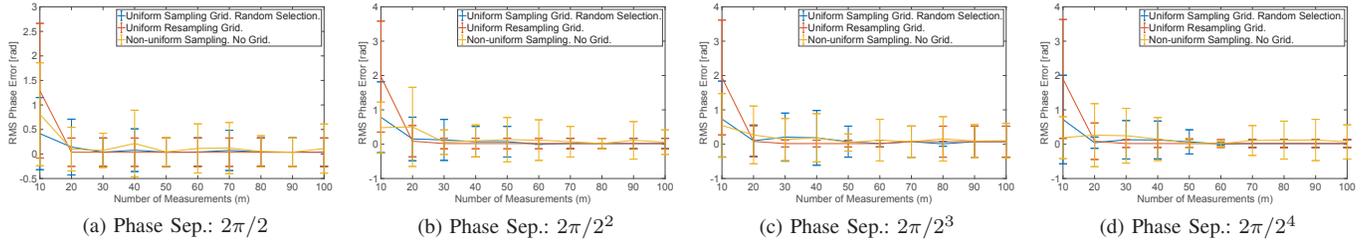
Fig. 3: Phase RMSE of nested RIAA for different separations between targets and different sampling schemes in frequency domain. Both separations and errors are given in phase domain, where 2π corresponds to an unambiguous range of $15\,\mathrm{m}$, due to the minimum frequency of $10\,\mathrm{MHz}$. Mean and standard deviations were obtained over 64 experimental runs with different scene responses. For each plot, all randomly-generated scene responses showed two reflective targets, restricted to be at the desired separation from one another. The three sampling schemes considered are random selection on a regular grid (blue line), uniform sampling using a regular grid of the desired size (red line) and pure nonuniform sampling, without grid (yellow line).

harmonics $\omega_i = i\omega_0, i = 1, \ldots, m$. Uniform sampling allows using very fast parametric spectral estimation methods, like those based on Prony's method. This was the approach adopted in [6], [16], [17], where the *matrix pencil method* [18] was used to tackle the very same problem. If enough uniform frequency samples are available, even a simple inverse FFT followed by peak detection may suffice to find the main cisoid components of the data.

A. Nonuniform Sampling

If we are to deal with arbitrary sets of sampling points $\{\omega_i\}_{i=1}^m$ methods relying on uniform sampling no longer apply and we have to resort to *nonuniform* spectral estimation methods. It is out of the scope of the paper to provide a review on nonuniform spectral estimation methods and we focus exclusively on Stoica's Iterative Adaptive Approach (IAA) [19], which has shown to provide superior results among comparable approaches. Note that the measurements delivered by the ToF hardware are not complex, but real and, depending on the ToF pixel, it may require two physical acquisition to obtain the complex measurements in (6), using different phase shifts. For this reason, the real-valued IAA (RIAA) [20] is specially appealing for our application. In this paper we adopt RIAA to estimate the unknown parameters from the real (or the imaginary) part of (6) measurements that are nonuniformly distributed in frequency domain. Both IAA and RIAA have shown to widely outperform simpler approaches, such as the (nonuniform) Fourier transform, the least-squares periodogram or the Schuster periodogram. The latter approaches suffer from massive leakage and exhibit multiple false peaks, while the former are able to yield a clean spectrogram with peaks at the right locations.

1) Nesting the RIAA method: In working with experimental data we have observed that RIAA fails to estimate some of the smallest spectral components in the signal. In such cases we observed that, despite a peak appears at the right spectral location, a more prominent false peak appearing close to the main component is wrongly selected instead. Nevertheless, the main component is always very accurately recovered. For this reason and due to the typically low number of interfering paths in the ToF case, we propose applying RIAA recursively in a *nested* manner, estimating only the parameters of the main sinusoidal component in the data at a time. After the main component is estimated by RIAA, it is removed from the data and the RIAA runs again on the residual in an OMP fashion. This process is repeated K times, that is, as many times as sinusoidal components are expected in the data. Summarizing:

1) Initialize residual: $\vec{r} = \vec{y} = [y_i]_{i=1}^m$.
2) Run RIAA on $\vec{r}$ and obtain the main sinusoidal component.
3) Subtract the main sinusoidal component from $\vec{r}$.
4) Return to step 2 until K iterations have taken place.

We have observed that the nested RIAA yields much more accurate estimates of the parameters in practice.

V. RESULTS

In this section we present some experimental results, both using simulated and real data from pixels of a ToF camera, which demonstrate the feasibility of solving the MPI problem from a set of nonuniform frequency measurements. The nested-RIAA approach in-

troduced in Section IV will be used as nonuniform spectral estimation method for estimating the unknown parameters $\{\Gamma_k, d_k\}_{k=1}^K$.

A. Results using real data

For the experiments in this section we use a subset of the dataset used in [5], [11], which consists of 77 complex frequency measurements between $10\,\mathrm{MHz}$ and $86\,\mathrm{MHz}$ (both included) with $1\,\mathrm{MHz}$ step, acquired with a ToF camera featuring a PMD 19k-S3 sensor. Provided that the samples are uniform, the matrix pencil method can be readily applied to them. Doing so the parameters are indeed correctly estimated. In order to evaluate the effect of dropping samples and the feasibility of the nonuniform sampling scheme, we consider different subsampling ratios in $(0, 1]$. For the nested-RIAA we randomly pick subsets of samples from the dataset. In order to allow for a fair comparison, for the matrix pencil method we select a subset containing the lowest-frequency measurements. Figure 2 shows the normalized RMSE of the reconstructed data using the parameters estimated by the two methods for three ToF pixels. The scene response functions contained three targets in all cases. Clearly, for the same number of samples, nonuniform sampling and nested-RIAA estimation is superior to classical uniform sampling and estimation via the matrix pencil method.

B. Results using synthetic data

In this section we evaluate the performance of nested-RIAA in distinguishing two targets at different distances between them. The fundamental frequency, defining the unambiguous range in the uniform sampling case, is set to $10\,\mathrm{MHz}$ and the bandwidth of the measurement system to $100\,\mathrm{MHz}$. Three different sampling schemes are considered. Figure 3 shows the results obtained for target separations of $2\pi/2^r$, $r \in \{1, \ldots, 4\}$, in phase domain for the fundamental frequency. The method exhibits good average performance for the range of target separations considered. All sampling schemes (see figure caption for details) seem to yield similar average performance for resolving two targets.

VI. CONCLUSION AND FUTURE WORK

In this paper we have tackled the problem of estimating a sparse scene response function from a set of nonuniform frequency measurements. The problem naturally arises in ToF imaging, where nonuniformity of the sampling relates to physical limitations. A thorough experimental evaluation with both synthetic and real data from a ToF camera provides a solid evidence of the feasibility of the approach. To the best of our knowledge this is the first work attempting to tackle the ToF MPI problem using nonuniform frequency sampling.

In future work we will consider a larger set of nonuniform spectral estimation methods to solve the problem of estimating sparse scene responses from nonuniform Fourier samples, including parametric and semi-parametric (sparsity-driven) methods. We will also consider further analysis of sampling bounds and study the applicability of nonuniform spectral estimation methods to cases where the sampling lattices are prefixed by the hardware. One further application case we will tackle in future work that is close to ToF imaging is fluorescence lifetime imaging microscopy (FLIM).

REFERENCES

[1] 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," *Nature Communications*, vol. 3, no. 1, jan 2012.

[2] C. Saunders, J. Murray-Bruce, and V. K. Goyal, "Computational periscopy with an ordinary digital camera," *Nature*, vol. 565, no. 7740, pp. 472–475, jan 2019.

[3] A. Kadambi, V. Taamazyan, S. Jayasuriya, and R. Raskar, "Frequency domain TOF: encoding object depth in modulation frequency," *CoRR*, vol. abs/1503.01804, 2015.

[4] A. Kirmani, A. Benedetti, and P. Chou, "Spumic: Simultaneous phase unwrapping and multipath interference cancellation in time-of-flight cameras using spectral methods," in *IEEE Intl. Cong. on Mult. and Expo (ICME)*, Jul. 2013, pp. 1–6.

[5] A. Bhandari, A. Kadambi, R. Whyte, C. Barsi, M. Feigin, A. Dorrington, and R. Raskar, "Resolving multipath interference in time-of-flight imaging via modulation frequency diversity and sparse regularization," *Opt. Lett.*, vol. 39, no. 6, pp. 1705–1708, Mar. 2014.

[6] A. Bhandari, M. Feigin, S. Izadi, C. Rhemann, M. Schmidt, and R. Raskar, "Resolving multipath interference in kinect: An inverse problem approach," in *SENSORS, 2014 IEEE*, Nov. 2014, pp. 614–617.

[7] A. A. Dorrington, J. P. Godbaz, M. J. Cree, A. D. Payne, and L. V. Streeter, "Separating true range measurements from multi-path and scattering interference in commercial range cameras," in *Proc. SPIE*, vol. 7864, 2011, pp. 786 404–786 404–10.

[8] J. P. Godbaz, M. J. Cree, and A. A. Dorrington, "Closed-form inverses for the mixed pixel/multipath interference problem in amcw lidar," in *Proc. SPIE*, vol. 8296, 2012, pp. 829 618–829 618–15.

[9] M. Heredia Conde, *Compressive Sensing for the Photonic Mixer Device - Fundamentals, Methods and Results*. Springer Vieweg, 2017.

[10] A. Bhandari, C. Barsi, and R. Raskar, "Blind and reference-free fluorescence lifetime estimation via consumer time-of-flight sensors," *Optica*, vol. 2, no. 11, p. 965, nov 2015.

[11] 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.*, vol. 33, no. 5, pp. 45–58, Sep. 2016.

[12] B. Buxbaum, R. Schwarte, T. Ringbeck, H.-G. Heinol, Z. Xu, J. Olk, W. Tai, Z. Zhang, and X. Luan, "A new approach in optical broadband communication systems: a high integrated optical phase locked loop based on a mixing and correlating sensor, the photonic mixer device (pmd)," in *Proceedings / OPTO 98, Internationaler Kongress und Fachausstellung für Optische Sensorik, Messtechnik und Elektronik, 18. - 20. Mai 1998, Kongresszentrum Erfurt*, 1998, pp. 59 – 64.

[13] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[14] E. J. Candès, "Compressive sampling," in *Proceedings oh the International Congress of Mathematicians*, Aug. 2006, pp. 1433–1452.

[15] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[16] M. Heredia Conde, T. Kerstein, B. Buxbaum, and O. Loffeld, "Fast multipath estimation for PMD sensors," in *5th International Workshop on Compressed Sensing Theory and its Applications to Radar, Sonar, and Remote Sensing (CoSeRa 2018)*, Siegen, Germany, Sep. 2018.

[17] A. Bhandari, A. M. Wallace, and R. Raskar, "Super-resolved time-of-flight sensing via FRI sampling theory," in *IEEE Intl. Conf. on Acoustics, Speech and Sig. Proc.*, Mar. 2016, pp. 4009–4013.

[18] Y. Hua and T. K. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 38, no. 5, pp. 814–824, May 1990.

[19] P. Stoica, J. Li, J. Ling, and Y. Cheng, "Missing data recovery via a nonparametric iterative adaptive approach," in *IEEE Intl. Conf. on Acoustics, Speech and Sig. Proc.*, Apr. 2009, pp. 3369–3372.

[20] P. Stoica, J. Li, and H. He, "Spectral analysis of nonuniformly sampled data: A new approach versus the periodogram," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 843–858, Mar. 2009.